UNIVERSITY OF HAWAII
LIBRARY
Dec 13 '61

Journal of MOLECULAR BIOLOGY

EDITOR-IN-CHIEF: J. C. KENDREW

Volume III, Number 5, October 1961

ACADEMIC PRESS LTD., LONDON
ACADEMIC PRESS INC., NEW YORK

JOURNAL OF MOLECULAR BIOLOGY

EDITOR-IN-CHIEF

J. C. KENDREW, Peterhouse, Cambridge, England

ASSISTANT EDITOR

S. BRENNER, M.R.C. Unit, Cavendish Laboratory, University of Cambridge, England.

EDITORIAL BOARD

P. DOTY, Department of Chemistry, Harvard University, Cambridge, Mass., U.S.A.
A. F. HUXLEY, Department of Physiology, University College, London, England.
M. MESELSON, Biological Laboratories, Harvard University, Cambridge, Mass., U.S.A.
R. L. SINSHEIMER, Division of Biology, California Institute of Technology, Pasadena, Cal., U.S.A.
J. D. WATSON, Biological Laboratories, Harvard University, Cambridge, Mass., U.S.A.
M. H. F. WILKINS, Physics Department, King's College, London, England.

ADVISORY BOARD

S. BENZER, Purdue University, U.S.A.
M. CALVIN, University of California, Berkeley, U.S.A.
F. H. C. CRICK, University of Cambridge, England.
J. F. DANIELLI, King's College, London, England.
A. ENGSTRÖM, Karolinska Institutet, Stockholm, Sweden.
A. D. HERSHEY, Carnegie Institute, Cold Spring Harbor, U.S.A.
F. JACOB, Pasteur Institute, Paris, France.
A. KORNBERG, Stanford University, Stanford, California, U.S.A.
C. LEVINTHAL, Massachusetts Institute of Technology, Cambridge, Mass., U.S.A.
S. E. LURIA, Massachusetts Institute of Technology, Cambridge, Mass., U.S.A.
D. MAZIA, University of California, Berkeley, U.S.A.
P. B. MEDAWAR, University College, London, England.
L. E. ORGEL, University of Cambridge, England.
M. F. PERUTZ, University of Cambridge, England.
R. D. PRESTON, University of Leeds, England.
G. N. RAMACHANDRAN, University of Madras, India.
A. RICH, Massachusetts Institute of Technology, Cambridge, Mass., U.S.A.
G. SCHRAMM, Max-Planck Institut, Tübingen, Germany.
F. S. SJØSTRAND, Karolinska Institutet, Stockholm, Sweden.
I. WATANABE, Institute for Virus Research, Kyoto University, Japan.
R. C. WILLIAMS, University of California, Berkeley, U.S.A.

Published by Academic Press Inc. (London) Limited

Volume III, 1961: 128*s.* 6*d.* ($18.00)

Subscription orders should be sent to Academic Press Inc. (London) Limited, 17 Old Queen Street, London, S.W.1, with the exception of those originating in the U.S.A., Canada, Central America, and South America; these should be sent to Academic Press Inc., 111 Fifth Avenue, New York 3, N.Y.

Copyright © 1961, by Academic Press Inc. (London) Limited

J. Mol. Biol. (1961) **3**, 483–506

The Molecular Structure of Collagen

ALEXANDER RICH

Biology Department, Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.

AND

F. H. C. CRICK

Medical Research Council Unit for Molecular Biology, Cambridge, England

(*Received 23 February 1961*)

This paper describes in detail our work on the structure of collagen which we have already outlined elsewhere (Rich & Crick, 1955). The main substance of the paper is:

(1) a demonstration that, given certain assumptions, only two basic types of structures are possible for collagen;

(2) detailed work on the coordinates and Fourier transforms of one of these models (collagen II), and a comparison between these predictions and the observed X-ray diffraction data.

1. Previous Work

The fact that collagen gives a distinct X-ray pattern has been known since the 1920's. The main features are a strong meridional spacing of 2·86 Å and an equatorial spacing of 11 to 15 Å which varies with humidity. We shall not discuss the earlier models (see the reviews by Bear, 1952, and Kendrew, 1954) which assumed that the structure consisted of sheets of chains. The first suggestion that the structure might be helical was made by Pauling & Corey (1951) who proposed a structure (now withdrawn) having three chains helically intertwined. The major advance, however, was the recognition of the correct non-integer screw axis by Cohen & Bear (1953), and by Cowan, North & Randall (1953). The latter workers obtained a greatly improved X-ray diffraction pattern by keeping the specimen stretched and moist during the X-ray exposure, and this has made it much easier to recognize the helical nature of the structure.

The next important advance was made by Ramachandran & Kartha. The two structures they proposed are not correct but they are of the correct type. Their first structure (Ramachandran & Kartha, 1954) consisted of a group of three parallel chains, each having a threefold screw axis. The three chains were themselves related by translations perpendicular to the fiber axis—or by a threefold screw axis; the two descriptions are equivalent. The chains were held together by hydrogen bonds between different chains.

In their second paper Ramachandran & Kartha (1955) modified this structure to conform to the correct non-integer screw. This they did by giving an additional slow

twist to their group of three chains so that each chain, in addition to following a (roughly) threefold screw axis, also gradually coiled round the other two. In short, they proposed a coiled-coil structure. This basic idea we believe to be correct, although the actual model they described is in our opinion stereochemically impossible. Their model and the two models proposed here (collagen I and II) are all closely related. They differ only in the way the three chains are "phased" with respect to one another, and in particular in *which* hydrogen bonds are used to hold the three chains together, These differences, however, are not trivial.

At about the same time, Cowan, McGavin & North (1955) reported that they had been working on similar lines; Ramachandran (1956) suggested that structure II is more likely to be correct than structure I; and Bear (1956) described an alternative method of systematic model building which leads to just the same two possible structures.

More recently we have published a short paper (Rich & Crick, 1958) listing the restrictions on the side-chains for both models and stating our preference for collagen II. Burge, Cowan & McGavin (1958), who also favour collagen II, have published a set of coordinates for it. There is thus a considerable measure of agreement among workers in the field (including Pauling, 1958) that collagen has the general type of structure described here, and that collagen II is the best structure. The only exceptions are Huggins (1957) who has published a proposal for a single-chain structure and Millionova & Andreeva (1958, 1959) who have suggested a double-chain structure. For reasons we give later in this paper we believe it almost certain that both of these structures are incorrect.

Two related structures must be briefly mentioned here. The first is poly-L-proline (Cowan & McGavin, 1955) and the other polyglycine II (Crick & Rich, 1955). The configuration of the polypeptide backbone in both these structures is related to that of the models proposed here. In addition, the hydrogen bonding system seen in the polyglycine II lattice has provided us with the stimulus for developing the alternative forms of the collagen molecule which have been described.

(a) *Chemical evidence*

It has been known for many years that collagen has an unusual amino acid composition. The analytical evidence has been summarized by Eastoe & Leach (1958) for vertebrate collagens, and by Watson (1958) for a small number of invertebrate collagens. For more recent data see Piez & Gross (1959). Almost one third of the residues are glycine, and there are considerable amounts of proline (usually 10 to 12%) and hydroxyproline (usually 6 to 10%); the latter is an imino acid found only in collagen and related proteins. Little was known about the sequence of the amino acids until the work of Schroeder and his colleagues (Schroeder, Kay, Legette, Honnen & Green, 1954) on partial hydrolysates of gelatin. They showed that the sequences Gly.Pro and Hypro.Gly were common, and suggested that a frequent sequence might be Gly.Pro.Hypro.Gly. This idea has been strongly supported by the work of Kroner, Tabroff & McGarr (1955). They, too, found that Gly.Pro and Hypro.Gly were common, Pro.Gly was very rare and Gly.Hypro missing, but in addition they identified the peptides Gly.Pro.Hypro and Gly.(Pro, Hypro).Gly. Both groups pointed out that the peptides identified did not fit the general sequence formula -P-G-R- (where P means proline or hydroxyproline, G means glycine and R any other residue) demanded by some of the structures proposed earlier.

Further work has been reported by Grassmann and his colleagues (Grassmann, Hannig, Endres & Riedel, 1956; Grassman, Hannig & Schleyer, 1960) on an analysis of the tryptic digest of collagen. They have isolated 114 peptide fragments and have reported on the composition of 51 of them. All but three of these have close to 33% glycine, showing that this amino acid is distributed quite evenly through the molecule. Pyrrolidine rings, however, are distributed unevenly, and occur in high concentrations in some peptides, and are virtually absent in others. For example, one peptide with 62 amino acids has 28 proline and hydroxyproline residues. This is clearly incompatible with a sequence G-P-R, but is consistent with the existence of sequences Gly.Pro.Hypro.Gly. Grassmann *et al.* (1960) have also presented evidence from end-group and chemical analysis which suggests that many of these tryptic digest peptides spontaneously aggregate together, usually in groups of three.

(b) *Physical chemical evidence*

It is possible to obtain a soluble form of collagen, as originally shown by Orekhovich, Tustnaovsky, Orekhovich & Plotnikova (1948). Physical chemical studies including light-scattering, osmometry, viscosity and flow birefringence have shown that the molecule in solution has the shape of a rather rigid rod of molecular weight 345,000, with a length of 3000 Å and a "diameter" of 13·6 Å (Boedtker & Doty, 1956; Doty & Nishihara, 1958). This is true both for soluble calf skin collagen as well as soluble ichthyocol from the carp swimbladder. Icthyocol undergoes a dramatic change when heated slightly, with a decrease in molecular weight. By preventing re-aggregation in the denatured product Boedtker & Doty (1956) have found that after denaturation the molecular weight of this soluble collagen falls by a factor of 2·5, and that the molecule is converted from a thin rigid rod into a more collapsed form, which has been assumed to be that of a random coil. The size of the molecule produced by this thermal denaturation is such that it can traverse the whole length of the collagen molecule. These data have been interpreted to support the idea that the soluble collagen is composed of three separate polypeptide chains, of unequal length, running side by side along the length of the molecule.

Another important line of evidence is the correlation between content of hydroxyproline and thermal stability, for collagens of widely different origins (Gustavson, 1953; Takahashi & Tanaka, 1953). Gustavson (1954), from a study of acetylated bovine collagen has suggested that hydroxyproline in collagen forms a hydrogen bond OH . . . . OC with the carbonyl group of the polypeptide backbone of a neighbouring chain. Subsequently a careful study has been carried out (Doty & Nishihara, 1958) relating the thermal denaturation temperature to the hydroxyproline content of soluble collagen. Using three types of collagen with differing hydroxyproline content (calf skin, icthyocol and cod skin) they have found that the thermal denaturation temperature of the soluble collagen molecule depends upon hydroxyproline content in a manner similar to the thermal shrinkage temperature. Thus, higher hydroxyproline contents appear to stabilize the molecule even when it is in the soluble form.

More recently these facts have been explained by Burge & Hynes (1959) and Piez (1960) in terms of a stabilization of the collagen molecule by *both* proline and hydroxyproline. Thermal shrinkage temperature is related to the total imino acid content, and may be associated with rotational constraints about the pyrrolidine ring. This does not, of course, rule out an additional contribution to stability by hydrogen bonding through the hydroxy group of hydroxyproline.

(c) *Infrared evidence*

The important part of the infrared evidence concerns the dichroism of the bands associated with the C=O and N—H groups of the polypeptide backbone (Ambrose & Elliott, 1951; Fraser, 1953; Sutherland, Tanner & Wood, 1954). The results show clearly that on the average the C=O and N—H bonds are more nearly perpendicular to the fibre axis than parallel to it. Moreover, the normals to the planes of the peptide groups are, on the average, in an intermediate orientation, and certainly not parallel to the fibre axis, as suggested by one earlier proposal (Crick, 1954).

It is also claimed (Badger & Pullin, 1954) that the frequencies of the infrared bands show that *cis* amide groups are not present in collagen to any appreciable extent. The argument is based on evidence from diketopiperazine; it would perhaps be desirable to have additional evidence before this important point can be regarded as completely established.

Recently, Beer, Sutherland, Tanner & Woods (1958) have made a quantitative evaluation of various structural proposals for collagen using the dichroism of several infrared bands. In this analysis, they used some of the transition moments obtained by Sandeman (1955) by polarized infrared studies on a crystal containing a peptide bond. Beer *et al.* concluded that the earlier proposal by Pauling & Corey (1951) was not compatible with the infrared data, but that the data were in agreement with the structure proposed by Ramachandran & Kartha (1955) as well as with the structure of collagen II. Polarized infrared studies have also been reported by Bradbury, Burge Randall & Wilkinson (1958), who also find good agreement between experimental results and the predictions obtained from the collagen II model.

2. Experimental Details

X-ray diffraction photographs were taken from rat tail tendons which were excised from freshly killed rats. Individual fibers of the tendon ($\sim$ 100 mμ diameter) were teased apart and mounted directly in a stretcher so that they could be photographed under constant tension. The fibers were elongated from 6 to 8% and then placed in a helium-filled X-ray diffraction camera which was maintained at a high (98%) relative humidity. The photographs were taken on a Hilger microfocus X-ray tube using copper Kα radiation.

Optical transforms were made using two different instruments. One of these (Elliott & Robertson, 1955) used a thin metal plate in which holes were drilled at atomic positions. In the other instrument (Wyckoff, Bear, Morgan & Carlstrom, 1957) an optical mask was made by attaching black dots to white paper and then carrying out serial photographic reduction to create the final optical image. In this system the area of the black dot was proportional to the number of electrons in the C, N or O atoms. Hydrogen atoms were ignored in making these masks. Since the collagen molecule has tenfold symmetry, one projection corresponds to views from ten different angles about the helix axis. To get better averaging about the fiber axis, in some cases two projections were taken, each 18° apart, and the diffraction images were then superimposed on each other to produce an averaging corresponding to 20 equally spaced views around the fiber axis.

Intensity estimates were made on the X-ray diffraction photographs using two different methods. A visual estimate of diffracted intensity was made using an intensity strip. Other sets of data were obtained using the Joyce Loebl microdensitometer with a wide entrance aperture to get a smooth distribution of intensity. For this purpose scanning traces were obtained traversing along individual layer lines as well as at right angles to the layer lines. In view of the difficulty in obtaining accurate intensity data from the blurred type of reflection obtained from collagen fibers, the several estimates were averaged.

3. The X-ray Diffraction Pattern

This is shown in Plate I. Recent descriptions have been given by Ramachandran & Ambady (1954), Bear (1955) and Cowan, North & Randall (1955). The salient features are:

(1) A strong meridional reflection, spacing 2·86 Å, which increases to about 3 Å on stretching.
(2) A strong equatorial reflection, which changes with humidity, going from about 10·5 Å for dry collagen to about 14 Å for wet collagen.
(3) A strong diffuse blackening on and near the equator at a spacing around 4·5 Å especially at low humidities.

The other strong reflections are near-meridional, with spacings of about 10 Å and 4 Å. The latter was at one time thought to be truly meridional, but the improved stretched X-ray pictures show it to be off-meridional.

The observed layer lines, which are somewhat diffuse, can be indexed using helical diffraction theory (Cochran, Crick & Vand, 1952) and give a helical screw axis having a translation of 2·86 Å and a rotation of approximately $108°/n$ where n refers to the n-fold *rotation* axis (parallel to the screw axis) which may or may not be present. (In our models $n = 1$.) This concise description includes the "net ambiguities" of Bear (1955). In addition, the molecule could have dyad axes perpendicular to the fiber axis.

Density considerations suggest that the number of amino acid residues in the structure per 2·86 Å in the fiber direction is probably 3, though 4 is possible and 2 cannot be completely ruled out. The most recent discussion is that of Bear (1956).

We therefore have the following alternatives:

Table 1

n (n-fold rotation axis parallel to fiber axis)	Number of residues in the asymmetric unit	Number of residues per 2·86 Å
4	1	4
3	1	3
2	1	2
2	2	4
2†	1	4
1†	1	2
1†	2	4
1	2	2
1	3	3
1	4	4

† means additional dyad axis perpendicular to the fiber axis.
The preferred alternative is underlined.

The first three alternatives can be eliminated with complete confidence on two grounds. First, since there is only one residue in the asymmetric unit it is a simple matter to build all possible models. It is soon found that no satisfactory model can be built using the usual criteria of bond angles, distance and van der Waals contacts. This was first pointed out by Cowan *et al.* (1953). Second, the diffraction pattern of

such models, even if they could be built, would not give large enough intensities close to the meridian, as explained by Cowan, McGavin & North (1955).

The fourth type of model set out in Table 1 includes the structure suggested by Crick (1954) which is certainly incorrect. Here again experience in model building and also with the diffraction pattern of such models (Bear, 1955) suggests that no structure of such a type is possible, though since there are two residues in the asymmetric unit this conclusion is less easy to arrive at.

In a similar way, the models with dyads perpendicular to the chain axis can be ruled out. Two of them have one residue in the asymmetric unit and can be ruled out by trying to build all alternatives. We can thus eliminate all the first six cases in Table 1 and can conclude that $n = 1$. Notice that so far we have not assumed anything about amino acid sequences, hydrogen bonding or even whether the residues are *cis* or *trans*.

The fact that $n = 1$ does not mean that collagen has only one chain. This can be seen from Fig. 1 (reproduced from Bear, 1955) which shows the "connection ambiguities". The diffraction pattern symmetry tells us that there are 10 units in 3 turns, with a translation of 2·86 Å per unit. Figure 1 is constructed by placing a cylinder

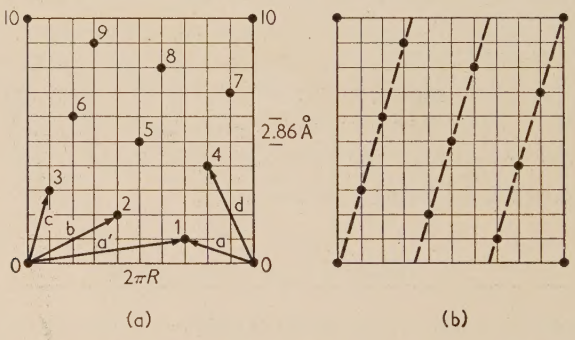


Fig. 1. The helix net of collagen, which has 10 units in three turns. (a) The units are numbered as dots on the figure. Arrows with letters represent various possible ways of connecting the units by covalent bonding.

(b) The helix net for collagen in which the dotted lines represent the polypeptide backbone connections for the three chain models discussed in this paper.

around the molecule on which dots represent the repeating units. The cylinder is then cut open parallel to the axis and laid flat. Lettered arrows in Fig. 1(a) represent possible ways of connecting the units. With minor reservations connections a or a′ will correspond to one chain, b to two, c to three and d to four. If there are only 3 residues per node, d is impossible, since a polypeptide chain cannot stretch so far.

It is difficult to proceed further without making some assumptions. The basic difficulty is that collagen is not a simple polymer, but a protein. Moreover, the 640 Å band-interband variations along the length of the molecule show that in some sense the molecule is not "uniform." In addition the X-ray pattern, even for the stretched material, is a somewhat diffuse one. It is therefore almost impossible to construct a logical "proof" of the structure, since many of the arguments are seriously weakened by the possible inhomogeneity of the material. The final proof of the collagen structure may have to await the production by the polymer chemist of "synthetic collagen" having a simplified amino acid sequence.

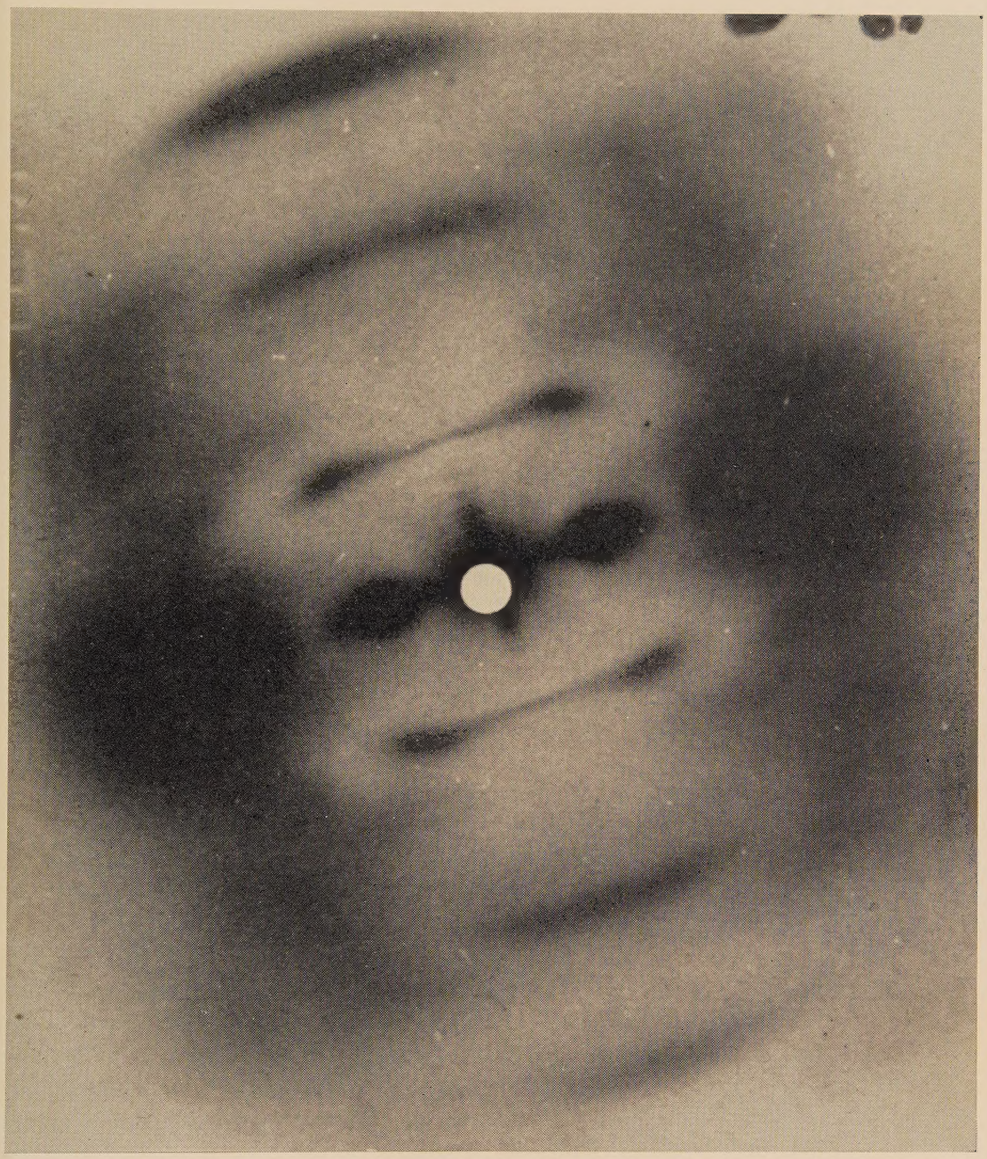


PLATE I. A high-angle X-ray diffraction pattern of rat tail tendon collagen. The fiber axis is about 20° from the vertical, and the specimen has been tilted to bring the 2·86 Å reflection at the top of the picture into the sphere of reflection.

[*To face page* 488

We shall therefore list a series of suggestive arguments which point towards the structures described here, realising that while this is a useful method for discovering structures it is a poor method for proving them.

(1) Good monochromatic X-ray pictures of collagen (T. Furnass, Jr., unpublished) give no hint of a second phase radically different from the main phase. It is probable, therefore, that the density argument leading to 3 residues per 2·86 Å is correct.

(2) The chemical data, taken together, suggest that the sequences Gly.Pro.Hypo.Gly, Gly.Pro.R_3.Gly, and Gly.R_2.Hypro.Gly are common ones, that is, that there are three residues (in succession on a single chain) in the asymmetric unit, and that every third residue is usually glycine.

(3) It has proved impossible to stretch collagen more than 10% of its length. The only exception occurs in certain electron microscope preparations (see Bear, 1952) but in our opinion this is probably spurious, and due to chains slipping past each other. We feel that this makes any single-chain model unlikely, and suggests that the polypeptide chains are fairly well extended.

(4) The physical-chemical data shows that collagen consists of more than one chain; the most likely number is three (Boedtker & Doty, 1956).

All these facts suggest that collagen consists of 3 chains (so that the method of connecting nodes is that shown in Fig. 1(b)) and that there are three residues per node.

We therefore built all possible models having these characteristics and assumed in addition:

(1) that all the peptide residues are *trans* (because of the infrared evidence) and approximately planar;

(2) that there is at least one set (out of three) of systematic hydrogen bonds. We feel that without such a set there would be nothing to hold the chains together in a regular manner;

(3) that the three chains are all parallel, rather than anti-parallel. This restriction is discussed more fully later;

(4) also that it should be possible to add pyrrolidine rings to the structure. This restricts to some extent the disposition of the peptide planes around the Cα carbon atoms. Apart from this we did not assume anything about the nature or sequence of the side-chains.

Our method was to assemble three lengths, side by side, of the polypeptide backbones using brass skeletal atomic models (scale 5 cm = 1Å) with all side-chains omitted. These chains were held loosely in approximately the arrangement necessary to give the observed screw axis. A particular set of hydrogen bonds was then selected. Starting from one particular NH group there are only a finite number of ways in which it can be joined to the C=O group of a neighbouring chain, since, if the C=O group selected is too far away, it becomes quite impossible to join up a systematic set of such hydrogen bonds. (By a systematic set we mean the set produced by the operation of the observed screw axis.) We actually attempted to make bonds to the six nearest C=O groups.

When one particular set had been selected and the relevant groups in the model had been joined with flexible wire, we attempted to manoeuvre the model so that the van der Waals contacts were acceptable. No configuration was rejected unless it grossly violated the permitted distances. For some sets it was possible to show by a simple logical dilemma ("either it fouls here or it fouls there") that no model could be

built. In other cases a large number of ways of packing had to be tried systematically before it became certain that none of them was acceptable. We feel reasonably sure that we have not overlooked a possible structure, though naturally it would be an advantage if other workers repeated our attempts.

Our conclusions were quite clear. The only structures that can be built satisfactorily are the two we have described as collagen I and collagen II or structures related to them. All other ways of making hydrogen bonds systematically are unacceptable. In addition we were able to show that no structure of this general type, but having *two* systematic sets of hydrogen bonds, can be built satisfactorily.

Within the rather narrow limits we have imposed, this model building may be said to be satisfactory. We should briefly note two other attempts at systematic model building on different assumptions. Cowan, North & Randall (1955) report that they have built "nearly one hundred models" all of which they considered unsatisfactory. However, the assumptions they made were perhaps not particularly well chosen, and as they give no systematic account of what they tried, the value of their work is difficult to assess.

Bear (1956) also has carried out systematic model building. He made the assumption that the repeated sequence Gly.Pro.Hypro must be possible (which we did not) but, in contrast to us, he made no initial assumptions about hydrogen bonds. He assumed, as we did, that the peptide bonds were *trans* and planar, that there might be 2, 3 or 4 chains (we assumed 3), and that the structure must fit the observed screw. He found seven possible configurations, but only two of them had satisfactory van der Waals contacts and hydrogen bonds. These were the structures we have called collagens I and II. Dr. Bear has informed us that this model building was done after he had heard that we had derived two structures but before he knew what they were, so that the two attempts may be considered independent. The fact that the same result has been arrived at from two rather different sets of assumptions is thus significant.

4. Alternative Models

Since our paper was published, a radically different model has been put forward by Huggins (1957). This model consists of one chain, with three residues in the asymmetric unit, and is made to fit the observed screw axis. Huggins objects to our models because they make too few systematic hydrogen bonds (one out of three) and his model attempts to make two out of three.

We think that Huggins's model is incorrect for the following reasons:

(1) It is difficult to build. We are not certain that it is stereochemically feasible.

(2) We regard *any* single chain model as unsatisfactory because we should expect such a structure to be extensible, at least under some conditions. Although extension has been reported in the electron microscope, we believe, as stated earlier, that this is spurious, and in any case the extensibility of Huggins's model is not enough to account for such a big stretch without subsidiary hypotheses, such as the super-folding he discusses.

(3) The model cannot accommodate the sequence Gly.Pro.Hypro without gross distortion. This is bound to be the case for any model that attempts to make two out of three hydrogen bonds systematically, since the position which the Cδ atom of the pyrrolidine ring most occupy will necessarily be too close to the neighbouring oxygen to which the nitrogen is normally hydrogen bonded.

(4) Its predicted diffraction pattern is an extremely bad fit with the X-ray picture. We have derived very approximate coordinates for Huggins's model and have calculated the diffraction pattern from them. The disagreement is particularly bad on the 3rd, 6th and 10th layer lines and small alterations of the structure will not alter this. This illustrates the point that not every structure built to the correct screw will give a satisfactory diffraction pattern.

An alternative suggestion for a collagen structure has been presented by Millionova & Andreeva (1958, 1959) who have considered a model with an alternating Gly.Pro sequence, since some of the tryptic peptides obtained by Grassmann *et al.* (1956) had a high concentration of glycine and imino acids. The molecule consists of two chains, hydrogen bonded together. The chains are related by a twofold axis parallel to the fiber axis. However, since each chain has 10 Gly.Pro units in three turns each with a translation of 2·86 Å, the symmetry is incorrect for the collagen diffraction pattern. Its predicted diffraction pattern would have all the odd layer lines absent (including the 3rd and 7th which are very intense). Accordingly, we do not believe that it can be a constituent of the collagen molecule, since, as cited above, there is no evidence for another structural component in collagen.

5. The Two Structures

We have already described these before in outline (Rich & Crick, 1955, 1958). It is not easy to grasp the arrangements of the atoms without studying a 3-dimensional model, since drawings tend to have a confused appearance. The easiest way to visualise them is to start from the (ideal) structure proposed for polyglycine II. The *left*-handed version ($P3_2$) must be chosen since poly-L-proline has a left-handed screw axis. A projection, perpendicular to the fiber axis, of two chains from this structure is shown in Fig. 2(a). It can be seen that each polypeptide chain has a left-handed threefold screw axis. To make this easier to grasp a schematic drawing of the same view is shown in Fig. 2(b), and further schematic views of a group of *three* chains from the polyglycine II lattice are shown in Figs. 2(c) and 2(d), since such a group of three chains can be selected in two different ways.

Imagine that the group of three chosen is that shown in Fig. 2(c). These three chains are then deformed so that they slowly twist round each other in a gradual *right*-handed helix, as shown schematically in Fig. 2(e). This gives the basic structure for collagen I. If the group of three chains had been that illustrated in Fig. 2(d) (and careful study will show that this group is not identical in its inter-relationships with that marked I) we should have arrived at collagen II.

In polyglycine II the screw axis which relates a residue in one chain to the next residue up in a neighbouring chain is a threefold screw axis. This screw axis has a rotation of 120° and a translation of 3·1 Å. The effect of deforming the three chains—as described above—to give collagen I or collagen II is to alter the screw to a rotation of 108° (approximately) and a translation of 2·86 Å in the direction of the fiber axis. It will be realised that this deformation is not very great.

There is a difference in the hydrogen bonding in the two models. In structure I the NH groups point anti-clockwise when viewed from the carboxyl ends of the chains, in structure II, clockwise. This is shown in Fig. 3, where we are looking down the axis of the molecule. The nomenclatures of different authors are shown in Table 2.

The two structures are also distinguished by the sequence of hydrogen bonds along any one of the three chains, as shown in Fig. 4.

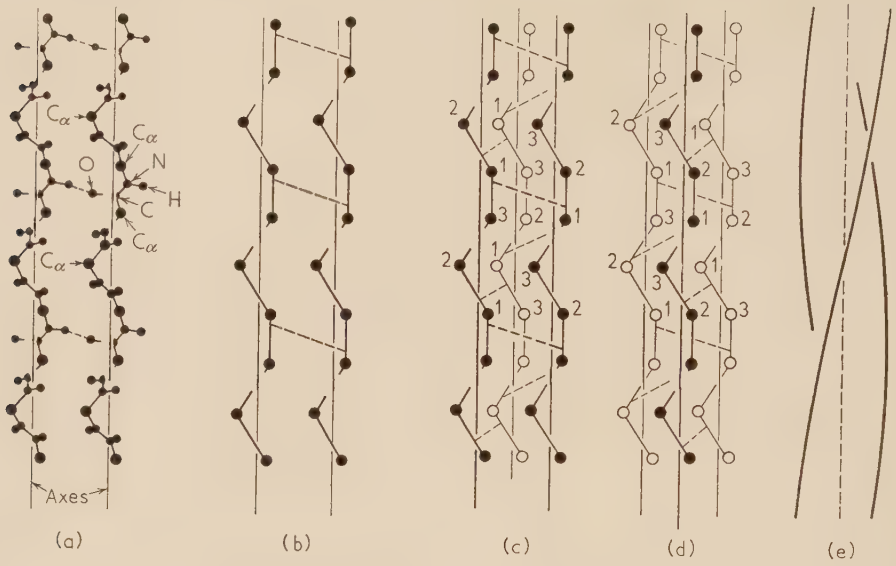


Fig. 2. Diagram illustrating the formation of the two collagen models from the polyglycine II lattice.

(a) Two strands of polyglycine with black dots representing various atoms and dashed lines as hydrogen bonds.

(b) Two chains of the polyglycine II lattice in which only the Cα atoms and hydrogen bonds are shown.

(c) The third chain shown with open circles lies behind the two shown in (b) to make a collagen I arrangement. The numbers indicate the phasing of the residues on the polypeptide chains.

(d) The third chain is added in front of the two in (b). The chain in front is shown by solid lines. This gives rise to the collagen II model.

(e) Solid lines represent the axes around which the polyglycine chains are coiled. These axes are gently coiled around each other in the collagen molecule.

Fig. 3. An axial projection of a segment of the two collagen models. The filled circle in the center represents the axis of the molecule. Three peptide groups are shown. The dashed line represents a hydrogen bond.

It will be clear from the comparison of Figs. 2 and 3 with Fig. 4 that within the group of three chains only one third of the possible hydrogen bonds are made (neglecting for the moment the effects of proline and hydroxyproline) and that they are made

in a different manner in the two structures. Given a model of collagen I, how would one change it into collagen II? The first step is to break the existing set of hydrogen bonds. Then each chain must be rotated *about its own* (*curved*) *axis* by about 60°. The NH group which previously made the hydrogen bond to one neighbouring chain (Fig. 3) is now attached to the carbonyl oxygen on the other neighbour. If all such bonds are made, we shall have gone from I to II. Thus the two structures differ in the way the two chains are "phased" with respect to one another.

TABLE 2

Rich & Crick (1955)	collagen I	collagen II
Cowan, McGavin & North (1955)	anti-clockwise	clockwise
Ramachandran (1956)	plus structure	minus structure

So far we have not discussed the position of possible side-chains. To do this we shall number the residues on one chain successively 1, 2, 3, 1, 2, 3, 1, 2, etc. The operation of the basic left-handed screw (108° and 2·86 Å) takes us from position 1 on one chain to position 1 on its neighbour. A second operation of the screw takes us to position 1 on the third chain, and a third operation brings us to position 1 on the original chain

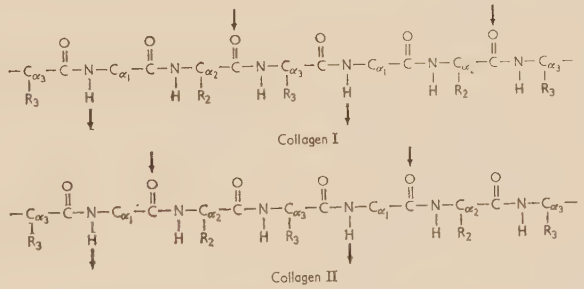


FIG. 4. A schematic diagram showing the polypeptide backbone and the systematic hydrogen bonds (arrows) which are formed in models of collagen I and collagen II. Although the amino group of glycine residues is involved in both of the hydrogen bonds, the carbonyl group of R_2 is involved in collagen I and the carbonyl group of glycine in collagen II.

but naturally three residues higher than the position 1 from which we started. Thus to go from position 1 on one chain to the next position 1 higher up on the same chain takes a left-handed screw of $3 \times 108° = 324°$ and $3 \times 2{\cdot}86 = 8{\cdot}58$ Å; that is, a right-handed screw of 36° and 8·58 Å. This can also be seen from the net diagram, Fig. 1(b). The operation described above for changing a model of collagen I into one of collagen II leaves the *numbering* of the residues unaltered.

It will be recalled that a common sequence in collagen appears to be Gly.Pro. Hypro.Gly. In both structures each backbone can, by itself, accommodate L-proline or L-hydroxyproline residues in positions 2 and 3 without difficulty. This is because the backbone arrangement is similar to that found in poly-L-proline by Cowan & McGavin (1955). If the mirror-image versions of the three polyglycine chains had been selected, then L-proline or L-hydroxyproline could not have been added to the backbone. It is this fact which fixes the "hand" of the screw axis of the collagen structure.

In our original paper (Rich & Crick, 1955) we stated that both collagen I and collagen II were stereochemically satisfactory. This is true of the backbones, but when side-chains are added the situation is more complicated. The addition of side-chains to collagen II leads to no difficulties as far as we can see. Moreover the glycine site is very close to the center of the structure and situated in such a way that no amino acid residue is possible there except glycine, even if the structure is deformed. That every third residue is glycine is suggested by the chemical evidence—for example every tripeptide (except one) which has been reported contains glycine. In addition, almost all of the tryptic peptides contain one third glycine. It will be interesting to see how much of the amino acid sequence follows this rule.

For collagen I, however, it is necessary to deform the simple arrangement of the backbones in order to add the C_β atom at position 3, since it would otherwise make too close a van der Waals contact with the peptide group of the neighboring chain. The structure is even more restrictive if hydroxyproline is added at position 3 in such a way that it makes a hydrogen bond from its OH group to the C=O of the neighboring peptide group. In fact it seems unlikely that three chains each having the repeating sequence Gly.Pro.Hypro could form structure I. However, only about one third of the residues in position 3 are hydroxyproline. It seems certain, however, that to build structure I to accommodate the side chains found in collagen it would have to be deformed from the regular arrangement. Moreover, while only glycine can be allowed in position 1 for an undeformed structure only a small deformation is required (as pointed out by Ramachandran, 1956) to permit other residues to go there. The restrictions imposed on the side chains are tabulated in Table 3.

TABLE 3

The possible positions of side-chains

Position	Collagen I		Collagen II
	undeformed	deformed	
1	glycine only	other residues may be possible: Pro or Hypro impossible	must be glycine
2	any residue including Pro and Hypro		any residue including Pro and Hypro
3	glycine only	any residue, including Pro and Hypro, except valine and *iso*leucine	any residue, including Pro and Hypro
bonding of the OH of Hypro in position 3	—	*can* make a hydrogen bond to the neighboring chain within the group of three	sticks out radially away from the structure and *cannot* make a hydrogen bond *within* the group of three chains

It is difficult enough to obtain accurate coordinates for an undeformed structure of this complexity; for deformed structures, presumably having a certain amount of strain, it becomes exceedingly difficult. However, to check the diffraction pattern only approximate coordinates are required. We have thus been forced to make a compromise and present very accurate sets of coordinates for collagen II only (Table 4), and one of our rough working sets of coordinates for collagen I (Table 5). The coordinates given are for a screw of 108° and 2·86 Å.

TABLE 4

Coordinates of collagen II

		x(Å)	y(Å)	z(Å)	ϕ(deg.)	r(Å)
Lower hydroxy-proline	C_α	3·43	0·00	0·00	0·0	3·43
	C	2·74	1·01	0·91	20·3	2·92
	O	3·21	2·16	1·07	33·9	3·87
	N	1·63	0·57	1·47	19·3	1·73
	H	1·28	− 0·35	1·31	− 15·3	1·33
Glycine	C_α	0·82	1·38	2·37	59·2	1·60
	C	1·64	1·88	3·56	48·9	2·49
	O	2·65	1·26	3·96	25·4	2·93
	N	1·19	2·99	4·11	68·3	3·22
	H	0·38	3·47	3·77	83·7	3·49
	Cα	1·84	3·63	5·26	63·1	4·07
Proline	C_β	1·15	4·97	5·48	77·0	5·10
	C_γ	− 0·20	4·76	4·82	92·4	4·76
	C_δ	0·00	3·78	3·68	90·0	3·78
	C	1·69	2·75	6·51	58·4	3·23
	O	0·66	2·08	6·70	72·4	2·18
	N	2·72	2·79	7·33	45·7	3·90
	H	3·53	3·34	7·15	43·4	4·86
Upper hydroxy-proline	C_α	2·78	2·02	8·58	36·0	3·43
	C_β	4·12	2·34	9·25	29·6	4·74
	C_γ	4·58	3·63	8·55	38·4	5·84
	O (H)	6·04	3·65	8·46	31·9	7·06
	C_δ	3·97	3·59	7·14	42·1	5·35

It should be remembered that the object of presenting coordinates to 0·01 Å is not because it is believed that the atoms are situated in the positions given to this accuracy but to show that at least a version of the structure can be built with acceptable bond distances and angles, even if the version presented is not quite the correct one.

We have listed for structure II the bond distances and angles calculated from our exact coordinates (Fig. 5(a)). It can be seen that these are in excellent agreement with accepted values; the covalent bond distances are within 0·01 or 0·02 Å, and the bond angles within 1° and 2° of the "canonical" values of Corey & Pauling (1953). All the van der Waals distances are satisfactory (Fig. 5(b)) as is the hydrogen bond length of 2·8 Å. The hydrogen bond is straight to within 1·5°. While the coordinates of Burge *et al.* (1958) for structure II are acceptable, our own set, which agrees closely with theirs, is the better one. We do not consider that either of the two positions suggested by them for water molecules are at all likely.

(a) *Optical transform studies*

A preliminary comparison of the diffraction patterns produced by the two collagen models was made by making optical transforms from which all side chains were excluded. The repeating sequence was thus Gly-R_2-R_3 where R_2 and R_3 are amino acids which have only β-carbon atoms but do not have the pyrrolidine rings. Plate II (a) and (b) show the optical transforms of collagen I and II without any side chains.

TABLE 5

Coordinates of collagen I

		x(Å)	y(Å)	z(Å)	ϕ (deg.)	r(Å)
	C	0·0	2·3	0·0	0	2·3
Lower	C_β	— 1·05	3·0	0·9	— 19	3·15
hydroxy-	C_γ	— 1·8	3·9	0·0	— 25	4·3
proline	C_δ	— 1·4	3·6	— 1·4	— 21	3·85
	N	0·05	3·15	— 1·25	1	3·15
	O (H)	— 3·2	3·7	0·15	— 41	4·9
	C	1·35	2·3	0·8	30	2·65
	O	1·9	2·4	1·1	29	3·9
	N	1·8	1·1	1·15	58	2·1
Glycine	C_α	2·95	1·0	2·15	71	3·1
	C	2·45	1·75	3·45	54	3·0
	O	1·25	1·7	3·7	36	2·1
	N	3·35	2·35	4·15	55	4·1
	C_β	4·35	4·05	5·6	47	5·95
	C_γ	5·5	3·55	4·6	57	6·55
Proline	C_δ	4·8	2·45	3·9	63	5·4
	C_α	3·15	3·15	5·35	45	4·45
	C	2·95	2·25	6·6	53	3·7
	O	3·7	1·25	6·8	71	3·9
Upper	N	1·9	2·5	7·35	37	3·15
hydroxy-	C_δ	1·0	3·7	7·2	15	3·85
proline	C_α	1·35	1·85	8·6	36	2·3

These each represent the superposition of two diffraction patterns of the molecule, 18° apart. Thus they are an averaging of 20 views of each molecule. It can be seen that the two diffraction patterns are grossly similar but there are several differences in detail. Both models predict a strong meridional tenth layer line which dominates the diffraction pattern, as well as vertical columns of high diffracting intensity on either side of the center beam extending through most of the layer lines. This feature is also shared by the diffraction pattern of collagen (Plate I). However, in these diffraction patterns, it can be seen that the pattern produced by collagen II is closer to what is observed experimentally than that of collagen I. This will be seen more clearly when we make a comparison involving an averaging of *all* views about the fiber axis. This can best be done by calculation.

(a)

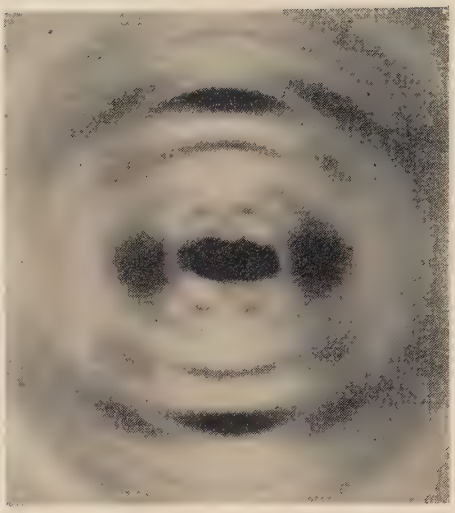
(b)

PLATE II. Optical transforms of collagen models with a sequence Gly-R_2-R_3, where R_2 and R_3 have only C_β atoms.

(a) Collagen I. (b) Collagen II.

[*To face page* 496

(a)



(b)

(c)

PLATE III. (a) An optical transform of one view of collagen II.
(b) The optical transform in (a) has been rotated $\pm$ 7° to illustrate the effect of disorder.
(c) An X-ray diffraction pattern of *unstretched* rat tail tendon collagen.
Photographs (a) through (c) are made on the same scale.

(b) *Effect of disorder*

In making a comparison between observed and predicted diffraction patterns, the optical transform method has serious disadvantages, especially in view of the limitations in approximating the atomic scattering factors and the difficulty of averaging different configurations. However, this method can be used very easily to illustrate the effect of disorder on the diffraction pattern. This is shown in Plate III. Plate III(a)

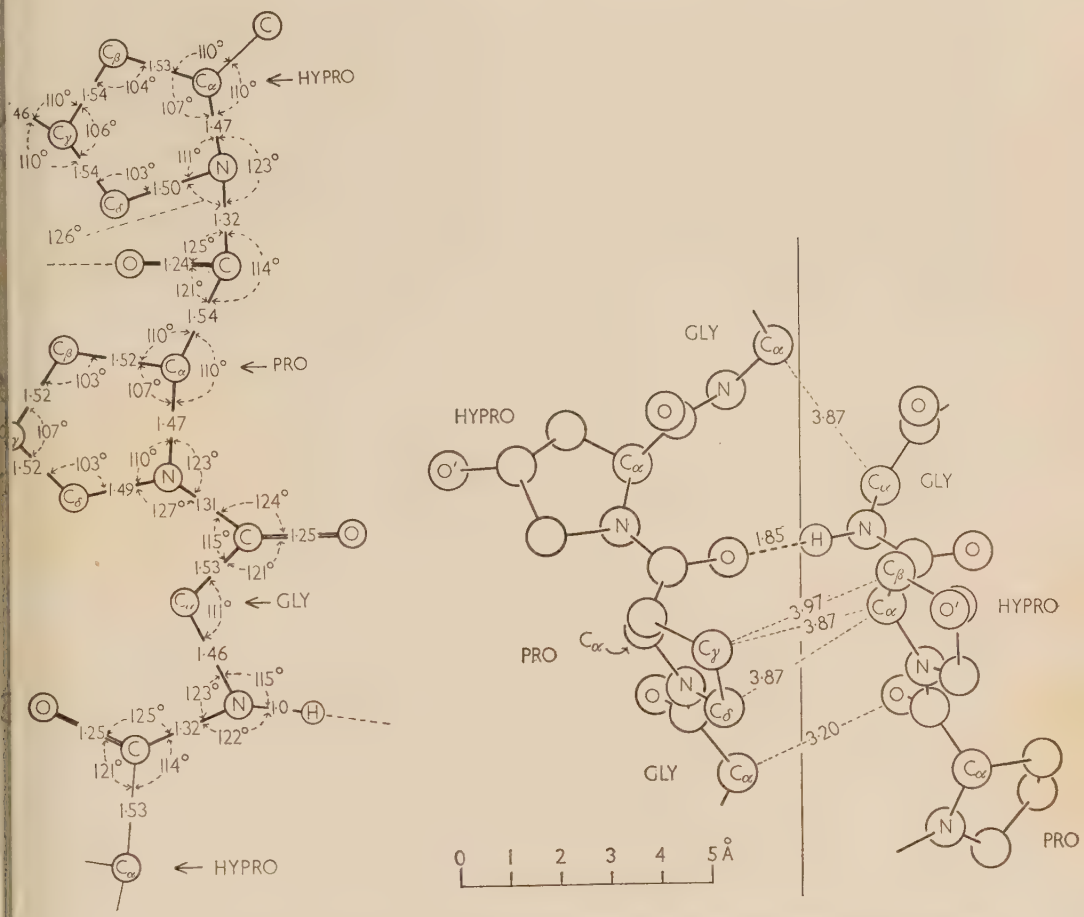


FIG. 5(a). A diagram (*not* to scale) of the bond angles and distances in collagen II for the repeating tripeptide sequence Gly.Pro.Hypro.

(b) A projection of two of the chains of collagen II to show the van der Waals contacts between chains. In this diagram it can be seen that the proline and hydroxyproline rings on adjacent chains lie more or less parallel to each other in van der Waals contact.

represents an optical diffraction pattern of *one* view of the collagen II molecule, in which β-carbon atoms are present on sites R_2 and R_3 and, in addition, an oxygen atom of water is hydrogen bonded to the same amino groups. The observed diffraction pattern of collagen is known to improve considerably on stretching. This is probably due to eliminating some disorder in the fiber. When unstretched the molecular axes are not all parallel to the fiber axis, but rather are distributed within some small angle.

This effect can be shown very clearly with the optical diffraction patterns. An oscillation of Plate III(a) about the central beam creates a diffraction pattern which is the superposition of many different orientations. Plate III(b) is a view of III(a) with an oscillation of $\pm 7°$. In the course of doing this, the detail in the diffraction pattern is eliminated and there remains only a strong meridional reflection at 2·86 Å (tenth layer line) and two other spots which represent the seventh and third layer line and diffuse scattering further away from the meridian. This is to be compared to a diffraction pattern produced by an *unstretched* rat rail tendon which is shown in Plate III(c). The meridional reflections in III(c) are of course weaker due to the geometry of X-ray diffraction in contrast to optical diffraction. Nevertheless, there are many similarities.

(c) *Measurement of observed intensity*

As pointed out above, the diffraction pattern of a collagen fiber depends very much on its mode of preparation. In the unstretched fiber, three main non-equatorial reflections can be seen corresponding to the third, seventh and tenth layer lines.

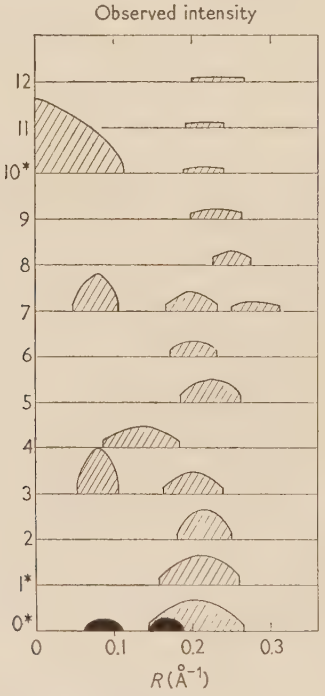


FIG. 6. The observed intensity of reflections from the X-ray diffraction pattern of rat tail tendon collagen. Intensity is plotted vertically for layer lines 0 to 12. The maximum of the dome-shaped cross-hatched areas represents the position of the observed reflection and the height is proportional to the maximum. The solid semi-circles on the equator represent very intense discrete reflections. *, intensity plotted at one half scale.

However, as a hydrated fiber is stretched, there is a gradual emergence of the other layer lines on the diffraction pattern. In the limit of extreme elongation, one can begin to see a separation of most layer lines. The third and seventh layer lines remain clearly separate and are very strong with their first maxima lying close to the meridian of the diagram. The fourth layer line has a broad diffuse peak lying somewhat closer

to the meridian than the remaining layer lines. The first and second layer lines are very broad and intense and are never completely resolved from the equator even on stretching the fiber to extreme limits. However, most of the other layer lines can be separated from each other. In view of the extreme diffuseness of the layer lines, it is very difficult to obtain accurate intensity data. In Fig. 6, we have plotted the observed intensity data on a reciprocal lattice grid. To emphasize the semi-quantitative character of the data, the reflections are indicated in the cross-hatched areas with a height which is proportional to the observed intensity. However, the cross-hatched region is arbitrarily taken as dome-shaped since it was felt that more detailed curves might give the illusion of presenting information which was more quantitative than that which was actually obtained. On the equator, the two black semi-circular areas represent the positions but not the intensities of the strong 12 to 14 Å and 6 Å reflections. The intensities listed on the equator and on layer lines 1 and 2 were obtained by making densitometer tracings along the layer line, even though the layer lines were not resolved from each other. For locating the intensity maxima, photographs were measured which had varying exposures to the X-ray beam. In this way, the peaks of the diffraction maxima could be located. The width of the diffraction peaks is indicated by the width of the cross-hatched areas and represents the width of the diffraction area down to an arbitrary low level of diffracting intensity. Because of the diffuseness the measurements in Fig. 6 cannot be regarded as having greater than 50% accuracy. For this reason no attempt was made to refine them by using polarization factors. In effect we have treated the whole diagram (except the equator) as the averaged transform of a single molecule. A similar attempt was made earlier by Bradbury *et al.* (1958) to measure the intensity of the equator and layer lines 3, 7 and 10. In Fig. 6 this attempt has been extended to all of the first twelve layer lines in the collagen diffraction pattern.

6. Calculated Diffraction Patterns

The considerable limitations of using optical transform methods for comparing the observed and predicted diffraction patterns can best be overcome by calculating diffraction patterns predicted by the various models. The diffraction pattern of collagen suggests that all orientations about the fiber axis of the molecule are found in any very small region of the fiber, since the layer lines have streaks rather than spots. Thus, we have used a computer program (Davies & Rich, 1959) which calculates the cylindrically averaged Fourier transform of a helical molecule. If we have a series of j atoms with cylindrical coordinates r_j, ϕ_j and z_j, the program computes the intensity

$$F_{\bar{\psi}}^2\left(R_1 \frac{l}{c}\right) = \sum_n (A_n^2 + B_n^2) \tag{1}$$

where

$$A_n = \sum_j f_j J_n (2\pi r_j R) \cos\left\{ n\left(\frac{\pi}{2} - \phi_j\right) + \frac{2\pi l z_j}{c} \right\} \tag{2}$$

and

$$B_n = \sum_j f_j J_n (2\pi r_j R) \sin\left\{ n\left(\frac{\pi}{2} - \phi_j\right) + \frac{2\pi l z_j}{c} \right\} \tag{3}$$

and where R and ψ are coordinates in reciprocal space, l = layer line number, c = repeat along the helix axis (= 28·6 Å) and J_n is a Bessel function of order n which satisfies the helical symmetry (Cochran *et al.* 1952). The three lowest values of J_n

were included in each layer line. The symbol $\bar{\psi}$ means we have averaged over all positions about the helix axis.

The basic difficulty in computing the diffraction pattern of a protein molecule such as collagen is the problem of the side chains. The coordinates which have been listed in Tables 4 and 5 are for a repeating polymer Gly.Pro.Hypro (Gly.-R-R_3). However, although the amino acid analysis of mammalian collagen shows that one third of the residues are glycine, the amount of proline is near 11% while hydroxyproline is 9% of the residues. This means that on the average only one third of the R_2 sites have proline residues on them, and likewise only about one third of the R_3 sites contain hydroxyproline. Two thirds of the R_2 and R_3 positions along the collagen molecule are thus occupied by all the other amino acids present in collagen. Our problem then is that of calculating the most appropriate Fourier transform to compare with the experimental data.

The computer program has sufficient flexibility to enable us to carry out different types of "averaging" for various molecular models. We can distinguish four different types of models for computing purposes. All of these are based on a repeating tripeptide sequence Gly-R_2-R_3.

(a) A model with Gly.Pro.Hypro using coordinates which are listed in Tables 4 and 5.

(b) A model with only the C_β atoms in R_2 and R_3. This means amino acids other than pyrrolidine rings or glycine are on those sites.

(c) A model with the C_β atoms in R_2 and R_3, but in addition an oxygen atom is hydrogen bonded to the N—H groups in R_2 and R_3. This is included in recognition of the fact that water molecules will bind onto these positions in the absence of proline or hydroxyproline residues. This model is a refined version of (b).

(d) Models such as the above with additional water molecules attached at various positions.

Calculations have been made using these types of models for collagens I and II. While there is not a great difference between the transforms calculated from models (b) and (c), a significant alteration is produced by model (a). The pyrrolidine rings markedly strengthen the calculated intensity on the third and seventh layer lines. For example, in collagen II, model (a) compared with model (b) produces a sixfold increase in predicted intensity for the near-meridional part of layer line 3, and about a twofold increase in layer line 7. The additional atoms of the pyrrolidine rings are thus largely in phase for the third layer line. In view of the magnitude of the effect we would predict that a comparative study of collagens with varying pyrrolidine content would show that the third layer line intensity varies directly with the amount of proline plus hydroxyproline.

Calculations have been carried out with various types of hydrated collagen models. One type of water molecule position is defined in (c) above, owing to the fixed position of the N—H bond. However, additional water molecules can be added at a variety of sites, and since they are in general attached to the collagen backbone by only *one* hydrogen bond, there is no easy way to decide on the particular positions to be used. The X-ray data are not good enough to be used to fix the position of these water molecules. Accordingly, we will not present the results of these investigations. Several sites for water molecules have been discussed by Bradbury *et al.* (1958).

We have carried out two different procedures for calculating the Fourier transforms of collagen. In one of these, we add together the transforms of model (a) (with

rings) and model (b) (without rings) in the ratio of 1 : 2. This corresponds to a physical arrangement in which the pyrrolidine rings are clumped together in the molecule. In another calculation, we have used model (a) but have assigned one third weight to the C_γ, C_δ and hydroxyl oxygen atoms on R_2 and R_3. This is the diffraction pattern produced by a random distribution of pyrrolidine rings throughout the molecule. While the predicted transforms obtained by both procedures are grossly similar, somewhat better agreement was found with the former calculation. This is interesting in view of the amino acid analysis of the tryptic digest peptides by Grassmann and his colleagues (1956, 1960) which showed that some peptides have large amounts of Pro + Hypro, while others have very little.

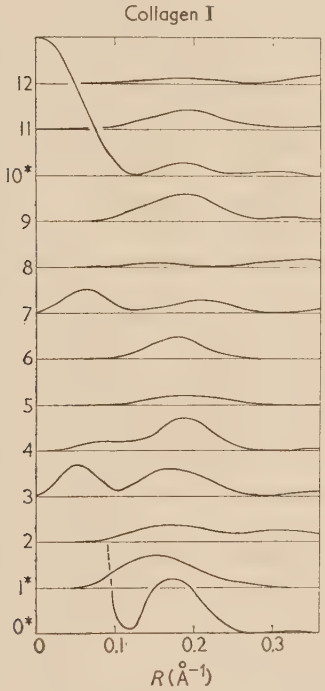


FIG. 7. The calculated transform of collagen I. The vertical distance represents the calculated intensity on each of the layer lines. *, intensity plotted at one half scale.

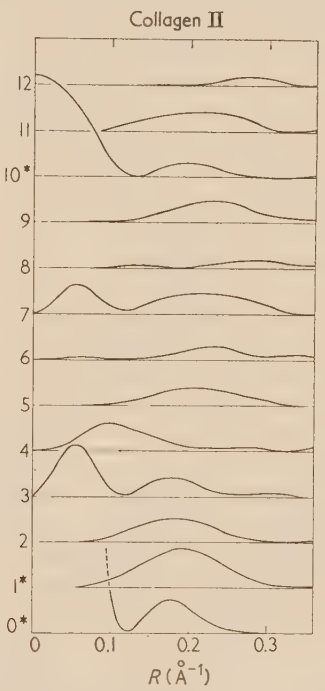


FIG. 8. The calculated transform of collagen II. Intensity is plotted vertically for the first twelve layer lines. *, intensity plotted at one half scale.

The intensities calculated in this way are shown in Figs. 7 and 8 for the two models. On the whole there is better agreement with the observed intensities for collagen II than for collagen I, although not too much weight should be put on this because of the neglect of water molecules and other factors, e.g. the diffraction pattern is for the stretched form but the calculations are for the unstretched dimensions. In collagen I (Fig. 7) the 10th layer line is too strong relative to the 3rd and 7th, while the 4th and 5th layer lines are relatively too weak. On the other hand, there is reasonable agreement between collagen II (Fig. 8) and the observed intensity. A slight discrepancy appears in the position of the first maximum of the 3rd layer line, which is found slightly further away from the meridian in the X-ray photograph. However, this 3rd layer line reflection frequently has a sharpness which suggests that, owing to

partial crystallinity, the transform may be sampled at that point. This may be associated with a type of helical or screw disorder due to the grooved character of the exterior of the molecule. Similar calculations have also been published by Bradbury *et al.* (1958) who also showed a qualitative agreement with collagen II.

7. Variants of the Basic Structures

It has been emphasized to us by Professor L. Pauling (1958) that if *no* assumptions whatsoever are made about side-chains two other simple structures are possible. These can be derived from the alternative model for polyglycine II (Crick & Rich, 1955) which has nitrogens near the axis. They can be derived also from collagens I and II *by rotating systematically all the peptide groups through 180°* about the C_α—C_α line. This breaks the hydrogen bonds, but since both the peptide groups involved are turned over, another hydrogen bond can be made in about the same place, as shown in Fig. 9(b).

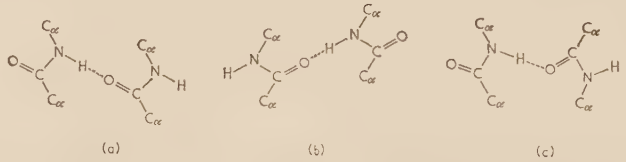


FIG. 9. Diagram illustrating some of the alterations that may be found in the collagen molecule.

(a) Represents a part of two strands of the collagen molecule which are hydrogen bonded together.

(b) The same two strands are hydrogen bonded together, but the planar peptide groups of the two strands have both been rotated 180° about the C_α—C_α positions. In this new configuration, another hydrogen bond forms.

(c) The chain on the right has been reversed so that the amino group is now lying below the carboxyl group unlike that of the earlier figures. It can be seen that in this position the altered chain is none the less still able to form a hydrogen bond.

These structures are not acceptable as models for collagen because it is quite impossible for them to accommodate pyrrolidine rings after the peptide groups have been turned over, since the C_β atoms no longer lie on the side of the peptide group which is required for ring formation. As far as we can see, these turned-over structures do not appear to have any compensating advantage.

However, it was suggested to us that we should systematically explore the effect of turning over *some* of the peptide planes of structures I and II, since pyrrolidine rings only occur in collagen in about one-third of the places in which structure II can accommodate them.

Another possible modification to the basic structures is to reverse one of the polypeptide chains so that the sequence of atoms in it is anti-parallel to that of the other two chains rather than parallel. We have reported (Crick & Rich, 1955) that this can be done in the polyglycine II lattice, and that the reversed chain can form satisfactory hydrogen bonds (as illustrated in Fig. 9(c)) and apart from the possible steric hindrance of the side-chains this can also be done for collagens I and II.

We have thus explored, by rough model building, all the possible variations of the structure which can be formed by combinations of these two operations, namely:

(1) the rotation of the plane of a peptide group by 180° about the C_α—C_α line (Fig. 9(b));

(2) the reversal of direction of one of the polypeptide chains (Fig. 9(c)).

One result of reversing the direction of the polypeptide chains is at first sight peculiar. Suppose one starts with structure II, reverses one chain, remakes the hydrogen bonds, and then repeats the process in turn for each of the two remaining chains one ends up with a structure which is "upside down"—the direction of all three chains has now been reversed—but the structure is now structure I. This can easily be understood, since the direction *in space* of the NH bond is more or less unchanged by this process (see Fig. 3), whereas the carboxyl ends of the chains is, after reversal, at the bottom of the structure (say) rather than at the top. Thus the direction of the NH bonds *viewed from the carboxyl end* has changed from clockwise to anti-clockwise, thus going from structure II to structure I.

In a structure with one or more chains reversed (but with none of its peptide planes turned over) each chain has a characteristic sequence of hydrogen bonds, either that found in structure I or that in structure II (see Fig. 4). Thus we may describe each *chain* as a type I chain or a type II chain. It is easy to see that there are only four basic structures (assuming for the moment that no peptide planes have been turned over), namely:

II II II
I II II
I I II
and I I I,

where, for example, I II II means that one chain has type I bonding and the other two have type II bonding.

We have tried turning over the peptide planes both singly and in groups for all these four basic structures, while preserving the basic number of hydrogen bonds. Turning over peptide planes converts pyrrolidine sites into sites which will only take other amino acids. We have systematically explored several variations of the four basic systems described above, but since we have no reason to believe they represent structural improvements over collagen II and have no evidence for their presence in collagen, we will not discuss them further.

8. Some Possible Side-chain Interactions for Structure II

We have examined, by rough model building, some of the possible interactions of side-chains *within* a group of three polypeptide chains for collagen II. We have for the moment restricted ourselves to considering hydrogen bonds between a side-chain and the polypeptide backbone, and have not considered those cases in which hydrogen bonds are formed between two side-chains.

We can find no satisfactory way of making intramolecular bonds between the backbone and either aspartic acid, asparagine, serine or threonine side-chains, mainly because these side-chains are too short; but, using glutamic acid, glutamine, lysine or arginine, one or more kinds of internal hydrogen bonds can be formed.

The details of these hydrogen bonds are given in Table 6. We have not permitted configurations in which an NH can make *no* hydrogen bonds (not even to water, that is) and we have discarded a few cases where the arrangement was very poor stereochemically. As criteria for the configuration of the side-chain, we have considered the van der Waals packing, the stagger of chemical bonds and the distance and angle of the hydrogen bonds.

It is interesting to note how few really satisfactory configurations there are. The three good ones make bonds between different polypeptide chains. No really good configuration has been found between a side-chain and its own backbone.

In studying the non-polar amino acids we find that the sequence -G-R_2-P- is hard to build if R_2 is valine or *iso*leucine, but we cannot be certain that it is impossible. Otherwise we have noticed no additional restrictions on the allowed amino acid sequence.

TABLE 6

Side-chain to backbone hydrogen bond interactions for structure II

Amino acid residue	Position of NH	chain	site	Position of C=O or OH	chain	site	Comment
Glutamic acid	backbone (no pyrrolidine ring)	C	2	sidechain	A	3	good
Glutamine	backbone	C	2	sidechain	A	3	good
	sidechain	A	3	backbone	A	3	poor
Lysine	sidechain	A	3	backbone	C	3	fair
	sidechain	A	2	backbone	A	3	fair
	sidechain	A	3	backbone	A	3	poor
Arginine	sidechain	A	3	backbone	C	3	good

The sites are numbered 1, 2 and 3 (site 1 is the glycine site). The three chains of collagen II are labelled A, B and C, going round the structure *anti*-clockwise looked at from the carboxyl ends of the chains (see Fig. 3). For descriptive convenience the side-chain has been considered to be attached to chain A, though naturally all three chains are equivalent.

We have not explored systematically hydrogen bonds between two adjacent groups of three polypeptide chains, but there is one possibility which should be mentioned, namely the formation of a pair of hydrogen bonds between an arginine on one group of three and two adjacent carbonyl groups on the backbone of one of the chains of an adjoining group of three. This can be done in a very satisfactory manner and may well be important. Moreover it is the only side-chain which can make *two* hydrogen bonds in this way. The maximum distance apart of the axes of the two groups is about 16 Å when the arginine residue is fully extended, but due to the flexibility of the hydrocarbon part of the side-chain, the bond can be made over a range of distances.

Arginine thus appears to be important for making hydrogen bonds to the polypeptide backbones both within a group of three chains and also between different groups. In this regard, it is worth noting that arginine and guanidine groups appear to play an important role in the formation of gelatin (Grabar & Morel, 1950; Berger, Kurtz & Noguchi, 1958; Janus, 1958) during which process it is quite probable that some of the collagen structure re-forms.

9. External Form of the Molecule

The shape of collagen II can best be seen with a model constructed from atoms with van der Waals radii. This is shown in Plate IV where the repeating sequence is Gly.Pro.Hypro. The individual chains are no longer separately visible; instead the molecule appears as an elongated cylinder with a deep groove helically wound around it. The groove, of course, arises from the glycine residues which occur at every third

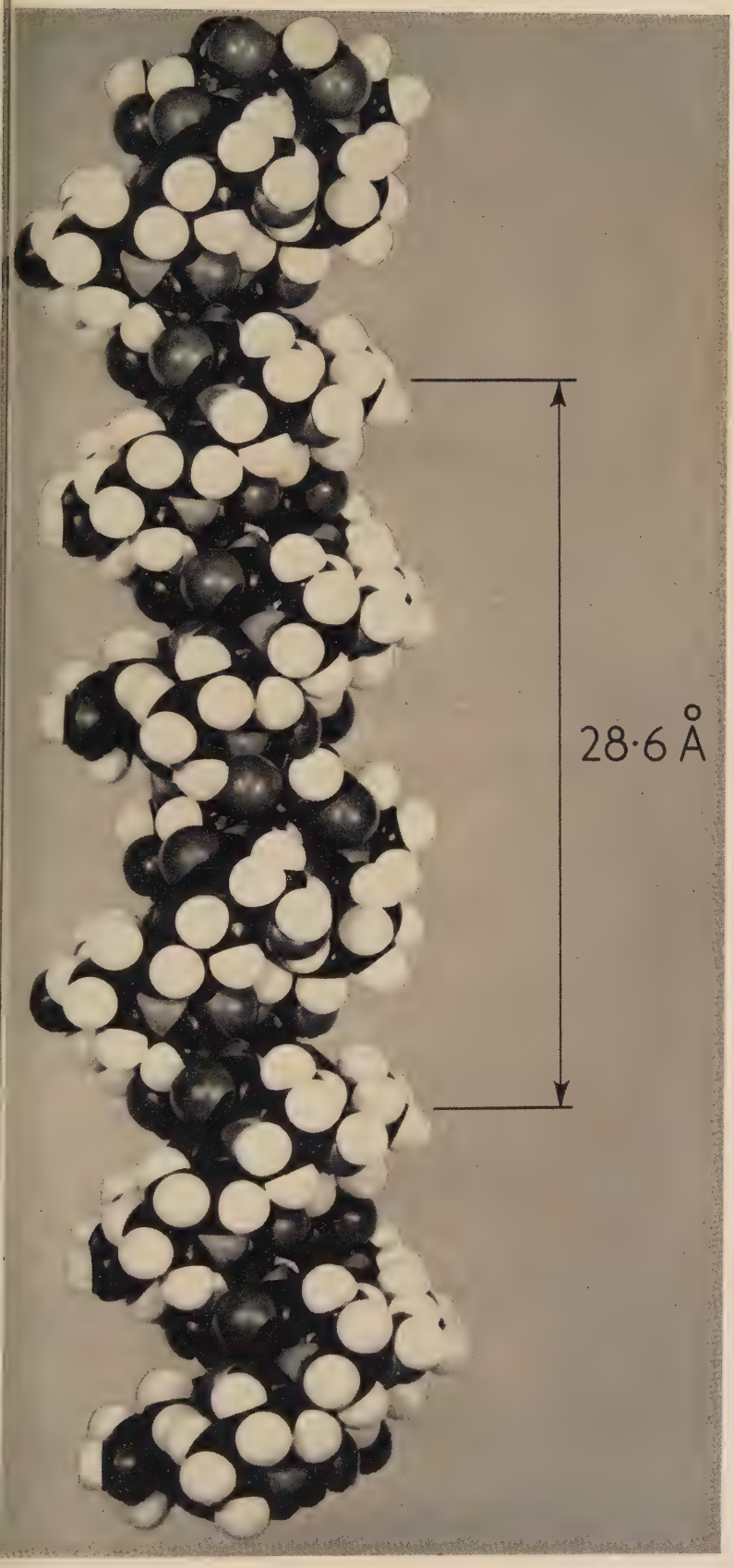


PLATE IV. Photograph of a model of collagen II made with space-filling models, which have been built with a repeating sequence Gly.Pro.Hypro. White atoms are hydrogen, black atoms carbon, light grey atoms are nitrogen (in planar peptide links) and dark grey atoms are oxygen. A repeat distance of 28·6 Å is marked.

[*To face page* 504

position along each chain and the corresponding lack of side chains. In collagen itself two thirds of the residues in the helical ridge would have amino acids other than proline and hydroxyproline. The longer side-chains would extend considerably beyond the pyrrolidine rings. When these molecules are laid down side by side, they are somewhat "screwed-together" since the ridge of one molecule tends to lie next to the groove of its neighbor. This may be important in stabilizing the molecule in the intact fiber.

10. Comparison of the Various Structures

There is little doubt that on aesthetic grounds one structure—collagen II—is superior to all the others. It is regular and very easy to build. It can accommodate the sequence Gly.Pro.Hypro at any point with no appreciable deformation, and its diffraction pattern is in good agreement with the X-ray picture of stretched collagen. It is thus quite likely that it is the major structural basis for the collagen molecule. However, it is not easy to prove that the collagen I or other modified arrangements are not found even though it is not easy to build these modifications in a regular way.

Experimentally the basic difficulty lies in the very poor quality of the diffraction pattern, especially for unstretched collagen, and for this reason we cannot rule out the existence of some of the variations we have listed, since their diffraction patterns would certainly be similar to those of the more regular structures. Lack of detailed knowledge of the amino acid sequence also makes it difficult to proceed further.

It is thus possible that collagen may be a mixture of structures. One of the striking features of the collagen diffraction pattern is the breadth of the layer lines. This suggests that the domain of regularity may be rather short and that this shortness may be due to backbone structural alterations. It is conceivable that both collagen I and II exist in the fiber, or some of the other variations described above. We have also considered as a possibility that the operation of stretching, which produces a greatly improved diffraction pattern, may also bring about some structural changes since it is accompanied by an increase in the fiber axis repeat. However, we cannot evaluate these possibilities at the present time.

The advantages of structures with one chain reversed is that a three-stranded structure can be produced from a single chain by super-folding of the type suggested by Huggins (1957), but whether this actually occurs we cannot tell.

Thus, in summary, while the basic plan on which collagen is built seems to have been discovered, much more work is required to reveal the precise details of the structure.

We should like to acknowledge the assistance of Dr. A. Elliott in the early stages of this work in obtaining optical transforms on his optical transform machine.

REFERENCES

Ambrose, E. J. & Elliott, A. (1951). *Proc. Roy. Soc.* A, **206**, 206.
Badger, R. M. & Pullin, A. D. E. (1954). *J. Chem. Phys.* **22**, 1142.
Bear, R. S. (1952). *Advanc. Prot. Chem.* **7**, 69.
Bear, R. S. (1955). *Symp. Soc. Exp. Biol.* **9**, 97.
Bear, R. S. (1956). *J. Biophys. Biochem. Cytol.* **2**, 363.
Beer, M., Sutherland, G. B. B. M., Tanner, K. N. & Woods, D. L. (1958). *Proc. Roy. Soc.* A, **249**, 147.
Berger, A., Kurtz, J. & Noguchi, J. (1958). In *Recent Advances in Gelatine and Glue Research*, ed. by G. Stainsby, p. 271. London: Pergamon Press.

Boedtker, H. & Doty, P. (1956). *J. Amer. Chem. Soc.* **78**, 4267.
Bradbury, E. M., Burge, R. E., Randall, J. T. & Wilkinson, G. R. (1958). *Disc. Faraday Soc.* **25**, 173.
Burge, R. E., Cowan, P. M. & McGavin, S. (1958). In *Recent Advances in Gelatine and Glue Research*, ed. by G. Stainsby, p. 25. London: Pergamon Press.
Burge, R. W. & Hynes, R. D. (1959). *J. Mol. Biol.* **1**, 155.
Cochran, W., Crick, F. H. C. & Vand, V. (1952). *Acta Cryst.* **5**, 581.
Cohen, C. & Bear, R. S. (1953). *J. Amer. Chem. Soc.* **75**, 2783.
Corey, R. B. & Pauling, L. (1953). *Proc. Roy. Soc.* B, **141**, 21.
Cowan, P. M. & McGavin, S. (1955). *Nature*, **176**, 501.
Cowan, P. M., McGavin, S. & North, A. T. C. (1955). *Nature*, **176**, 1062.
Cowan, P. M., North, A. T. C. & Randall, J. T. (1953). In *The Nature and Structure of Collagen*, ed. by J. T. Randall, p. 241. London: Butterworths.
Cowan, P. M., North, A. T. C. & Randall, J. T. (1955). In *Symp. Soc. Exp. Biol.* **9**, 115.
Crick, F. H. C. (1954). *J. Chem. Phys.* **22**, 347.
Crick, F. H. C. & Rich, A. (1955). *Nature*, **176**, 780.
Davies, D. R. & Rich, A. (1959). *Acta Cryst.* **12**, 97.
Doty, P. M. & Nishihara, T. (1958). In *Recent Advances in Gelatine and Glue Research*, ed. by G. Stainsby, p. 92. London: Pergamon Press.
Eastoe, J. E. & Leach, A. A. (1958). In *Recent Advances in Gelatine and Glue Research*, ed. by G. Stainsby, p. 173. London: Pergamon Press.
Elliott, A. & Robertson, P. (1955). *Acta Cryst.* **8**, 736.
Fraser, R. B. D. (1953). *J. Chem. Phys.* **21**, 1511.
Grabar, P. & Morel, J. (1950). *Bull. Soc. Chim. Biol., Paris*, **32**, 643.
Grassmann, W., Hannig, K., Endres, H. & Riedel, A. (1956). *Hoppe-Seyl. Z.* **306**, 123.
Grassmann, W., Hannig, K. & Schleyer, M. (1960). *Hoppe-Seyl. Z.* **322**, 71.
Gustavson, K. H. (1953). *Svensk. Kem. Tidskr.* **65**, 70.
Gustavson, K. H. (1954). *Acta Chem. Scand.* **8**, 1298.
Huggins, M. (1957). *Proc. Nat. Acad. Sci., Wash.* **43**, 209.
Janus, J. W. (1958). In *Recent Advances in Gelatine and Glue Research*, ed. by G. Stainsby, p. 214. London: Pergamon Press.
Kendrew, J. C. (1954). In *The Proteins*, ed. by H. Neurath & K. Bailey, Vol. II, p. 845. New York: Academic Press.
Kroner, T. D., Tabroff, W. & McGarr, J. J. (1955). *J. Amer. Chem. Soc.* **77**, 3356.
Millionova, M. I. & Andreeva, N. S. (1958). *Biofisika* (Eng. trans.), **3**, 259.
Millionova, M. I. & Andreeva, N. S. (1959). *Biofisika* (Eng. trans.), **4**, 138.
Orekhovich, W. N., Tustnaovsky, A. A., Orekhovich, K. D. & Plotnikova, N. J. (1948). *Biokhimiya*, **13**, 55.
Pauling, L. (1958). In *Recent Advances in Gelatine and Glue Research*, ed. by G. Stainsby, p. 11. London: Pergamon Press.
Pauling, L. & Corey, R. B. (1951). *Proc. Nat. Acad. Sci., Wash.* **37**, 272.
Piez, K. A. (1960). *J. Amer. Chem. Soc.* **82**, 247.
Piez, K. A. & Gross, J. (1959). *Biochim. biophys. Acta*, **34**, 24.
Ramachandran, G. N. (1956). *Nature*, **177**, 710.
Ramachandran, G. N. & Ambady, G. K. (1954). *Current Science*, **23**, 349.
Ramachandran, G. N. & Kartha, G. (1954). *Nature*, **174**, 269.
Ramachandran, G. N. & Kartha, G. (1955). *Nature*, **176**, 593.
Rich, A. & Crick, F. H. C. (1955). *Nature*, **176**, 915.
Rich, A. & Crick F. H. C. (1958). In *Recent Advances in Gelatine and Glue Research*, ed. by G. Stainsby, p. 20. London: Pergamon Press.
Sandeman, I. (1955). *Proc. Roy. Soc.* A, **232**, 105.
Schroeder, W. A., Kay, L. M., Legette, J., Honnen, L. & Green, F. C. (1954). *J. Amer. Chem. Soc.* **76**, 3556.
Sutherland, G. B. B., Tanner, K. N. & Wood, D. L. (1954). *J. Chem. Phys.* **22**, 1621.
Takahashi, T. & Tanaka, T. (1953). *Bull. Jap. Soc. Fish.* **19**, 603.
Watson, M. R. (1958). In *Recent Advances in Gelatine and Glue Research*, ed. by G. Stainsby, p. 179. London: Pergamon Press.
Wyckoff, H. W., Bear, R. S., Morgan, R. S. & Carlstrom (1957). *J. Opt. Soc. Amer.* **47**, 1061.

J. Mol. Biol. (1961) **3**, 507–519

Regularity in the Chain Configuration of D,L-copoly-γ-Benzyl Glutamate †

Akiyoshi Wada

Department of Chemistry, Faculty of Science, Ochanomizu University, Bunkyo-ku, Tokyo, Japan

(*Received 2 March 1961*)

The work described below leads to the following conclusions as to the structure of the D,L-copoly-γ-benzyl glutamate polymerized from a mixture of L- and D-N-carboxy anhydride (NCA) by initiation with strong bases.

1. Only one screw sense helix exists in the range $f_D = 0$ to 0·25.

2. Regularity in the orientation of polar groups in the backbone chain is maintained up to $f_D = 0{\cdot}35$.

3. The side-chain of the α-helix appears to have a regular orientation, as does its backbone chain.‡

4. In the right-handed α-helix, the electric dipole moment of the side-chain is opposed to that in the backbone chain for L-residues, and in the same direction in D-residues.

5. The electric dipole moment of the backbone chain of the α-helix is about 6·5 Debye units per residue in formamide-saturated ethylene dichloride.

6. Some disorder appears when more than $f_D = 0{\cdot}35$ of antipode is incorporated in the α-helix. This seems to originate in the unfolded parts at the ends of the chain or in a few disordered points in the molecule.

7. Even in the copolymer polymerized from an equimolar mixture of L- and D-NCA there is good reason to believe that the major portion of the substance is composed of helices which may be a mixture of right- and left-handed helices of L-residue-rich and D-residue-rich polymers, respectively.

Some of the above conclusions are derived by assuming rigidity in the side-chain. If the flexibility due to the internal rotational freedom allowed in carbon-carbon bonds in the side chain is taken into consideration, an alternative model is also possible:

1′. In an α-helix of pure L-polymer the electric dipole moment of the side-chain is forced in an opposite direction to that in the backbone by dipole-dipole interaction between the side-chain and backbone.

2′. When a D-residue is introduced, partial melting in the order of the side-chain occurs owing to steric hindrance, and the apparent residual moment will be increased by the decrease in the negative contribution of the side-chain.

3′. The contribution of the side-chain moment would not simply be offset in this case, and the moment of the backbone cannot be deduced from the above-mentioned extrapolation. But, since the melting in the order of the side-chain may be completed as the composition approaches that of the mesopolymer, the residual moment of the pure backbone of the α-helix is still believed to be 6 to 7 Debye units.

Although either of these models is plausible at the present stage, it is already evident that the side-chain in the α-helix, whether flexible or rigid, takes an

† This paper is No. VI in the series *Dielectric Properties of Polypeptide Solutions*. For the previous paper of this series see Wada, 1960.

‡ The same conclusion about the regularity of the side-chain has also been reached by the study of polarized infrared spectra by Blout & Miyazawa (1961), and by Tsuboi (1961).

ordered form and it can be assumed that the whole side-chain plays a role in determining the stability of the α-helix and its screw sense. The latter is supported by the evidence that a left-handed helix is stable in the case of poly-β-benzyl aspartate as reported by Bradbury, Downie, Elliott & Hanby (1960).

Finally, it must be mentioned that the above conclusions indicate that even in a complicated group of side-chains in a protein, there will be a fixed pattern rather than random orientation and that this is important in giving specificity to the character of the protein.

1. Introduction

The established fact that a poly-α-amino acid consisting of one kind of enantiomorphic residue forms a helical structure corresponding to one screw sense is attributable to the side-chain/side-chain interaction, which favors the one rather than the other sense. For poly-γ-benzyl glutamate (PBG) which is at present one of the most intensively studied polypeptides, a helix composed of L-residues is found to have a right-handed and one composed of D-residues a left-handed screw sense.

The major factor determining the screw sense is the interaction between the whole of the side-chain rather than that between the β-carbons only. Although in principle internal rotation is allowed in a carbon-carbon bond, it may be hindered to some extent in the α-helix because the side-chains are arranged so compactly in this configuration. If this is the case, the side-chains are likely to assume a regular orientation as does the backbone.

Accordingly, D,L-copoly-γ-benzyl glutamates of various D:L compositions were selected to study the chain regularity and the effect of incorporation of D-residues into an α-helix composed of L-residues. Optical rotatory dispersion measurement offers a unique method of estimating the helix content, and the dielectric method gives information on the orientation of polar groups in the backbone or the side-chain.

Several problems relating to the configuration of the D,L-copolymer have been examined using these two methods supplemented by consideration of polymerization kinetics.

2. Experimental

(a) *Materials*

D,L-Copolymer. Poly-γ-benzyl glutamate was selected in the present work because it is readily soluble in various solvents and because of the extensive data already available. Optically isomeric N-carboxy anhydrides (L-NCA and D-NCA) were synthesized from the corresponding glutamic acid † by the Blout-Karlson method (Blout & Karlson, 1956). Polymerization was carried out by dissolving the NCA of the required composition of D- and L-isomers in dry dioxane with sodium methoxide as an initiator. The concentration of the NCA for polymerization was approximately 5 g/100 ml. and the anhydride-initiator ratio was 100 in each case. After 3 days, when the extent of reaction was above 85%, the reaction mixture was poured into cold ethanol with vigorous stirring. The precipitate obtained was washed with ethanol and ethyl ether and dried *in vacuo*. The molecular weights of the samples thus obtained were estimated by measuring intrinsic viscosity in dichloroacetic acid (DCA) using the intrinsic viscosity/molecular weight relationship of Doty, Bradbury & Holtzer, 1956. In spite of an equivalence in the anhydride-initiator ratio in a series of polymerization reactions, a decrease in molecular weight with the fraction of D-isomer was found, as reported by Idelson & Blout (1958), and Downie, Elliott, Hanby & Malcolm (1957).

† The author wishes to thank the Ajinomoto Co. Tokyo, for supplying pure D-glutamic acid.

Information on the samples used is given in Table 1. In this Table, the fraction of isomer denotes the fraction of NCA used for polymerization. From measurements of the optical rotation of the samples in DCA, in which the polymer becomes a random coil and form rotation disappears, this fraction is found to be equivalent to the actual composition of the polymer. The further question whether the polymer obtained consists of homogeneous copolymers in which D- and L-residues are mixed with a given ratio in one molecular chain, or whether it is just a mixture of two kinds of polymers consisting of D- and L-isomers, respectively, is a major object of the present study and will be discussed later.

TABLE 1

Compositions and molecular weights of D,L-*copoly-γ-benzyl glutamate samples* †

Polymer no.	Fraction of D- and L- isomer		Mol. wt.	Intrinsic viscosity	
	f_L	f_D		DCA	EDC-FA
912	0·935	0·065	115,000	0·695	1·60
913	0·819	0·181	74,000	0·475	0·80
914	0·708	0·292	46,000	0·320	0·42
915	0·657	0·343	36,400	0·260	0·40
916	0·597	0·403	35,000	0·245	0·33
917	0·573	0·427	33,500	0·237	0·28
918	0·546	0·454	37,000	0·257	0·24
919	0·530	0·470	35,000	0·245	0·28
920	0·500	0·500	33,000	0·235	0·26
931	0·506	0·494	31,000	0·225	0·24
932	0·452	0·548	30,000	0·220	0·24

† Initiator: NaOMe, $A/I = 100$.

Solvents. Dioxane for the polymerization reaction was refluxed for 24 hr with sodium and fractionally distilled from sodium immediately before use (b.p. 101·0°C). Ethylene dichloride (EDC) and chloroform (CF) used for measurements of optical rotatory dispersion, dielectric constant and viscosity were dried over $CaCl_2$ and fractionally distilled (b.p. 84°C and 61°C, respectively). Formamide (FA) and dichloroacetic acid (DCA) were vacuum distilled.

(b) *Optical rotatory dispersion*

The measurements of optical rotation were made in the wavelength range 340 mμ to 700 mμ with the Hitachi photoelectric polarimeter with Beckman-type monochromator, using a 10 cm liquid tube. In the concentration range 0·8 to 1·0 g/100 ml. of PBG, EDC was used as the solvent.

(c) *Dielectric dispersion*

The dielectric constant was measured by the Schering bridge in the frequency range 150 to 40,000 c/s. A platinum concentric cylinder calibrated with pure benzene and EDC was used as a liquid cell in the oil bath, with the temperature controlled at 25 ± 0·05°C. Measurements were made in the concentration range 0·3 to 1·0 g/100 ml. and extrapolated to zero concentration. To avoid the molecular association previously reported (Doty *et al.*, 1956) (especially in chloroform) formamide-saturated chloroform (CF-FA) and formamide-saturated EDC (EDC-FA) were used as solvent systems. Details of the dielectric apparatus, measurement and the treatment of the data have been previously described (Wada, 1960*a*).

3. Results

(a) *Optical rotatory dispersion*

The dispersion curves obtained have been replotted for future convenience in a modified plot according to the procedure of Moffitt & Yang (1956), Yang & Doty (1957), Doty & Lundberg (1957*a*) and Imahori (1960) and representative plots are shown in Fig. 1. By assuming that contributions to the optical rotation are made only by

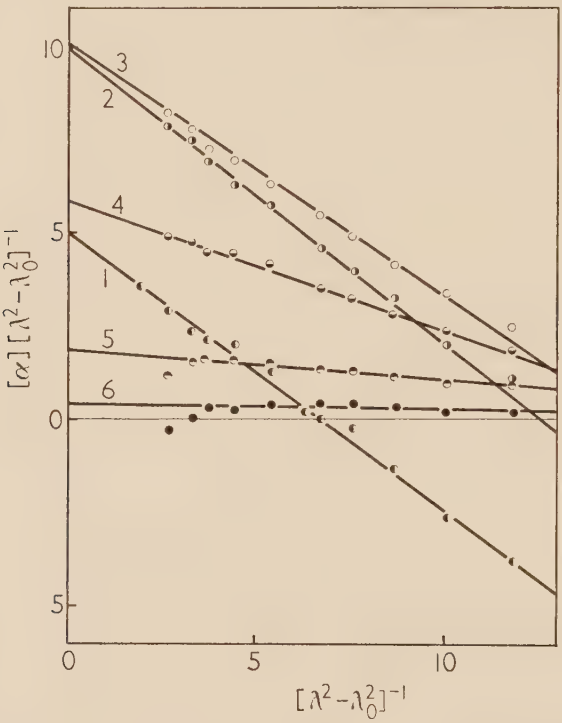


Fig. 1. Modified plots of the optical rotatory dispersion of D,L-copolymers. Fractions of D-residue; (1) 0·00, (2) 0·181, (3) 0·292, (4) 0·403, (5) 0·454, (6) 0·500.

the right-handed and left-handed helices and random coil, and that no other chain configuration exists in the present case, the specific rotation at the wavelength λ can be written in terms of the fraction of L- and D-residues, f_L and f_D respectively, and that of right- and left-handed helices, H_r and H_l, respectively, as follows:

$$[\alpha]_\lambda = \frac{100\,(n^2+2)}{3M_0}[m]_\lambda \tag{1}$$

$$[m]_\lambda = \{a_0^R(f_L - f_D) + a_0^H(H_r - H_l)\}\frac{\lambda_0^2}{\lambda^2-\lambda_0^2} + b_0^H(H_r - H_l)\frac{\lambda_0^4}{(\lambda^2-\lambda_0^2)^2} \tag{2}$$

where a_0^R denotes a constant, characteristic of constituent residues, that is, the amino acid residue, and a_0^H and b_0^H those for the helical configuration, n the refractive index of the solvent, and M_0 the residual weight. For the right-handed α-helix of L-PBG, the following constants have been used : $a_0^R = -540$, $a_0^H = 680$ and $b_0^H = -600$. The critical wavelength λ_0 was taken as 212 mμ.

In the modified plot of equation (2), $[\alpha]_\lambda(\lambda^2 - \lambda_0^2)$ *vs.* $(\lambda^2 - \lambda_0^2)^{-1}$, the intercept and slope are

$$I = \frac{100\ (n^2 + 2)}{3M_0}\{a_0^R(f_L - f_D) + a_0^H(H_r - H_l)\}\lambda_0^2 \quad (3)$$

$$S = \frac{100\ (n^2 + 2)}{3M_0}\{b_0^H(H_r - H_l)\}\lambda_0^4 \quad (4)$$

Consequently, only the difference between H_r and H_l can be independently obtained from both the intercept I and slope S of the modified plot. The values of $(H_r - H_l)$ thus obtained are plotted in Fig. 2, where it is shown that the left-handed helix does

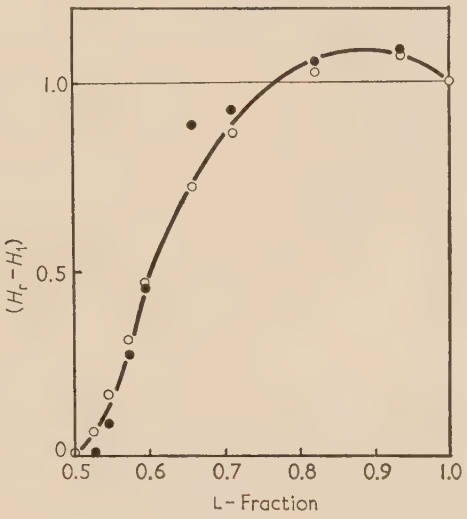


FIG. 2. Differences between the fraction of right-handed and left-handed helices. Calculated from intercept (○) and slope (●) of the modified plot.

not appear until 25% of D-residue is incorporated.† This result shows a well-known feature of such a copolymer, that is, the minor D-residues do not form left-handed helices but are incorporated in right-handed helices with the preponderant L-residues. Good agreement between the value obtained from I and S may suggest that the other regular forms such as the β-form do not exist (Wada, Tsuboi & Konishi, 1961) though in the meso-polymer they cannot be ruled out. The curves obtained for the different series of D,L-copolymers initiated by sodium methoxide and n-hexylamine coincide. This differs somewhat from the results obtained by Downie *et al.* (1957).

Although only the difference between the constants of the right- and left-handed helices, but not the absolute values, can be obtained by the above optical rotation method, the results of the dielectric study will be complementary and give information on chain regularity from a different angle, since the electric dipole moment given by this method is a function of the orientation of the polar groups in a molecule and is entirely independent of the helical sense.

† In Fig. 2 the values of $(H_r - H_l)$ exceed 1·0 in the range $f_D = 0{\cdot}0$ to 0·2. This apparently unusual feature suggests that some additional effect must be considered in equation (2), although it seems to be sufficient in the present discussion.

(b) *Dielectric dispersion*

The D,L-copolymer gives a typical dielectric dispersion curve in EDC-FA solution, as the representative plot in Fig. 3 shows, though the critical frequency of polymers of small molecular weight is beyond the range of the present apparatus. The dielectric constants are extrapolated to zero frequency to obtain the static dielectric constant from which the specific polarization P_2^0 is obtained. Then P_2^0 is extrapolated to zero

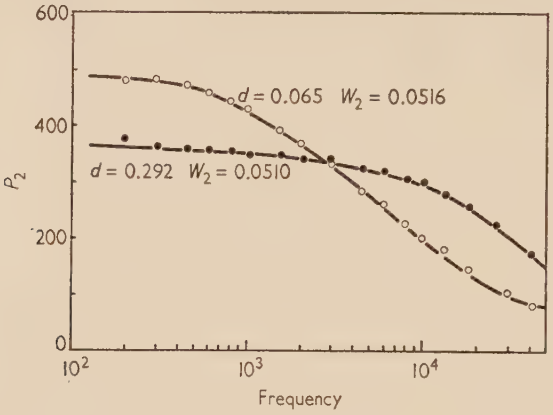


FIG. 3. Representative dielectric dispersion curves of D,L-copolymer of different D-fraction in EDC-FA.

concentration to give P_2^{00}, from which the electric dipole moment of the solute molecule is calculated. All of these procedures are explained in the previous report (Wada, 1960*a*). The measured P_2^{00} and the electric dipole moment of the D,L-copolymer are listed in Table 2 with other dielectric data.

TABLE 2

Dielectric results on the D,L-*copoly-γ-benzyl glutamate in* EDC-FA

Polymer no.	Specific polarization † P_2^{00} (cm³)	Dipole moment (Debye)	Residual moment (Debye)	Critical frequency
912	540	1775	3·34	12,000
913	490	1340	3·97	33,000
914	450	1010	4·80	90,000 ~ 130,000 ‡
915	440	890	5·37	140,000 ~ 250,000 ‡
916	300	720	4·50	—
917	230	610	4·44	—
918	200	740	4·49	—
919	260	665	4·15	—
920	270	660	4·38	—
931	270	640	4·50	—
932	290	645	4·70	—

† Specific electronic polarization is calculated to be 0·26 cm³ for PBG and is neglected in the present calculation.

‡ Estimated by the method described by Wada (1960*a*).

Although both the specific polarization P_2^{00} and the critical frequency ν_0 depend upon the fraction of the D-residue in the copolymer, conclusions cannot be drawn from the direct comparison of these values since molecular weight is also dependent on the D-fraction in the present samples as shown in Table 1. Since both P_2^{00} and ν_0 (the former being a function of the orientation of polar groups and the latter of the hydrodynamic shape of the molecule) can be given as a function of the degree of polymerization, the relation obtained by eliminating the degree of polymerization (shown by the bold line in Fig. 4) will be satisfied by perfect α-helices of any molecular weight (Wada, 1960*b*). A sharp deviation from this theoretical P/ν plot for the perfect α-helix has been observed in the case of D,L-copolymer when the D-fraction

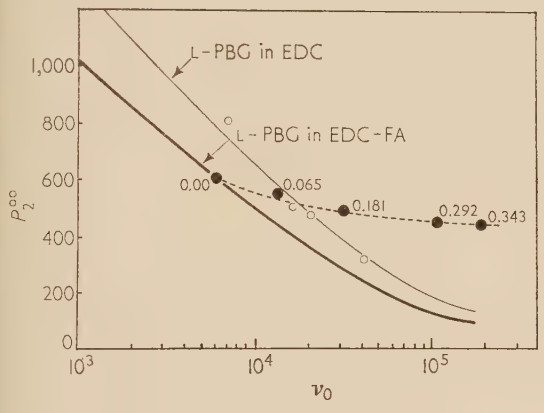


FIG. 4. Relation between specific polarization and logarithm of critical frequency for the perfect α-helix in EDC (open circles), and in EDC-FA. Filled circles represent the D,L-copolymers of different D-fraction.

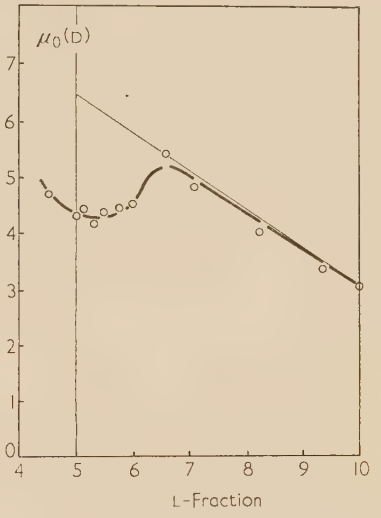


FIG. 5. Plot of μ_0, the electric dipole moment per residue, against the L-fraction. Narrow line of straight extrapolation gives a value 6·5 D at $f_D = 0{\cdot}5$.

increases, as shown in the same figure. By comparing the relation between P_2^{00} and the degree of polymerization, with the relation between ν_0 and the degree of polymerization, both for the D,L-copolymer and for the α-helix (Wada, 1960*a*), it is seen that the deviation from the α-helix of the pure L-polymer is reflected only in P_2^{00} even when the fraction of D- is small. This result shows that in advance of the change in hydrodynamic shape and of the helical structure, an alteration takes place in the orientation of polar groups owing to the incorporation of the D-residues into the polymer.

By calculating the electric dipole moment per amino acid residue, μ_0, from the molecular dipole moment and from the degree of polymerization, and plotting it against the fraction of L-, one observes an interesting feature, that μ_0 increases with the incorporation of the D-residue up to the $f_D = 0{\cdot}35$ range and then falls abruptly. (See Fig. 5.)

Before trying to elucidate this dielectric behavior in terms of the structure of the D,L-copolymer, let us consider the kinetics of copolymerization in a simplified model in order to form a general idea of the arrangement of the D- and L-residues on the polymer chain.

4. Simplified Model for the Polymerization Kinetics

From the abundant experimental evidence, the polymerization process of NCA may be considered as an addition polymerization in which the NCA molecules are added one by one to the N-terminal end of the polymer chain (Idelson & Blout, 1958; Lundberg & Doty, 1957). However, in the case of the copolymerization reaction of the optically isomeric NCA any theoretical treatment is complicated by the following considerations.

In the present case of the helical polymer, the effects due to residues other than nearest neighbors must be considered as well as the rate constants which appear in the usual copolymerization reaction such as k_{LL}, k_{LD}, k_{DL} and k_{DD}. Thus, the rate of addition will depend on the residue one turn before the chain terminus, that is, four residues away in the case of the α-helix; in this case we shall have 2^5 rate constants to consider. Even though the number is reduced by half by the symmetrical relation between the D- and L-residues, there are still too many to include in a numerical calculation. Consequently, the present discussion must necessarily be limited to a simple and empirical model though some ambiguity will still remain.

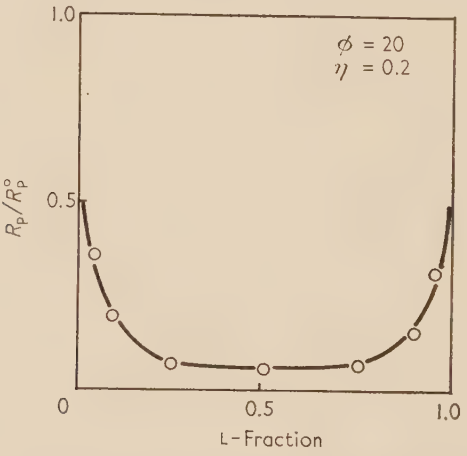


FIG. 6. Dependence of the rate of D,L-copolymerization upon the L-fraction. Rigid line is calculated by equation (5); $\eta = 0{\cdot}2$ and $\phi = 20$. Experimental data used were obtained by Idelson & Blout (1958).

First, let us try to explain the effect of the incorporation of D-residues in decreasing the rate of polymerization by considering the termination process. In the present process the slow cross reaction, such as ~~~ D+L or ~~~ L+D, seems to give opportunity for the termination. If this is so, the process may be similar to mono-radical termination and the rate of polymerization can be written in principle as follows: (Flory, 1953):

$$Rp = \frac{(2f_L^2 - 2f_L + 1) + 2\eta \,.\, f_L\,(1 - f_L)}{(2f_L^2 - 2f_L + 1) + 2\phi \,.\, f_L\,(1 - f_L)} \qquad (5)$$

$$\eta = k'_{DL}/k'_{LL}, \qquad \phi = k'_{tDL}/k'_{tLL},$$

where k'_{LL} $(= k'_{DD})$ and k'_{DL} $(= k'_{LD})$ are apparent rate constants and subscript t denotes termination reaction. By the use of the set of parameters in equation (5),

$\phi = 20$ and $\eta = 0{\cdot}2$, a fairly close agreement with the experimental results of Idelson & Blout (1958) can be obtained as shown in Fig. 6.

Unfortunately, because of the complex nature of copolymerization kinetics, no theoretical significance can be given to constant k' and the above discussion is a somewhat strained interpretation. However, the result obtained by Lundberg & Doty (1957) that $k_{LL}(= k_{DD})$ is about five times larger than that of the cross reaction, k_{DL} $(= k_{LD})$, seems to be consistent with the value of $\eta = 0{\cdot}2$, and it may not be meaningless to continue the discussion along this line on the basis that the agreement is not an accidental one. Using this constant η, the ratio in the D:L composition at the initial stage of polymerization, f'_L/f'_D, is given by the following equation:

$$\frac{f'_L}{f'_D} = \frac{f_L^2 + \eta(f_L - f_L^2)}{(1-f_L)(\eta f_L - f_L + 1)} \tag{6}$$

It is clear from this formula that a selective reaction occurs between the L-chain and L-residues and the D-chain and D-residues, and this fact will provide the basis of the following discussion.

Next, let us consider the growing chain in the copolymerization process, using several rules provided by empirically obtained facts:

(a) The rate of addition of a residue to a chain with a terminal residue of the same sense is about five times larger than the rate of addition to its antipode (Lundberg & Doty, 1957).

(b) An addition to the helical end of the favorable screw sense, that is right-handed for the L-residue and left-handed for the D-residue in the case of PBG, is also about five times faster than addition to the unfolded end (Doty & Lundberg, 1956, 1957*b*).

(c) An addition to the helical end of an unfavorable screw sense is possible (Lundberg & Doty, 1957) but will not occur so frequently.

(d) The helical configuration can be stabilized in the chain when more than four of the same kind of isomeric residue have an opportunity to link in sequence (Doty & Lundberg, 1956, 1957*b*).

(e) Although the helical sense would not be reversed by a few D-residues incorporated into the right-handed helix of the L-residue, reversal will occur when more than four D-residues are added in sequence to it (Lundberg & Doty, 1957).

On the basis of these empirical rules, copolymerization will proceed in the following manner: at first, the rate of initiation from L- and D-isomers is proportional to their ratio. These activated L- and D-residues will then tend to select the same kind of residues to be added, though such a selection may not be too marked. In this step the chain will grow without helix formation. However, once the chain forms a helical configuration of either screw sense as the result of a suitable sequence, such as LLLL or DDDD, the selection becomes increasingly positive and a helical molecule of a single screw sense will then be produced, since a reversal of the helical sense will not occur very often except in the region of an equimolar mixture of D- and L-residues.

The above situation strongly suggests that the major part of the D,L-copolymer will be a mixture of right-handed helices of L-residue-rich polymer and left-handed helices of D-residue-rich polymer, and not a homogeneous D,L-copolymer, although if the selection is less definite, several right- and left-handed helical parts will co-exist in a molecular chain, with random coil between them.

The probability of helix formation, that is the probability of getting m L-residues (or D-residues) linked together in sequence, is a function of the fraction of the L- and

D-residues. In the case of a positive selection, by putting the total fraction of the helix as H_0, the fraction of the right- and left-handed helix, H_r and H_l respectively, can be expressed in a simple form:

$$H_r = \{f_L^{m-1}/(f_L^{m-1} + f_D^{m-1})\}H_0 \quad (7a)$$

$$H_l = \{f_D^{m-1}/(f_L^{m-1} + f_D^{m-1})\}H_0 \quad (7b)$$

The value of $(H_r - H_l)$ empirically obtained from the optical rotatory dispersion seems to fit best the calculated curve for $m = 4$ in equation (7), as shown in Fig. 7. This is quite consistent with the fact that the necessary minimum number of residues to form a nucleus for a helical region is about four, as has been shown by studies on polymerization kinetics (Doty & Lundberg, 1956, 1957*b*).

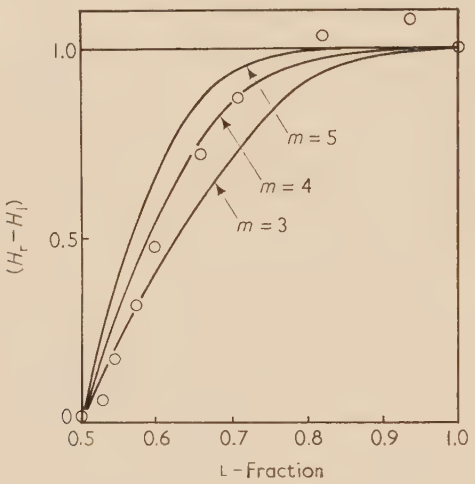


FIG. 7. Theoretical plot of $(H_r - H_l)$ against L-fraction for several values of m. Open circles obtained from the intercept of the modified plot.

Although the agreement between the experimental data and the calculated curve is seemingly sufficient, $(H_r - H_l)$ is observed to become a little smaller than the calculated value as it approaches $f_D = 0{\cdot}5$. Although the deviation is trivial, this might be a manifestation of the formation of something other than an α-helix, perhaps a random coil because sequences unfavorable for a helical configuration, such as DLDLDL, will increase in this range.

5. Discussion

We have now two sets of information concerning chain regularity, that is, the helical structure (from the optical rotatory dispersion) and the orientation of the polar groups in the molecule. From the former results it has been shown that the D,L-copolymer retains a perfect helical conformation of a single screw sense up to $f_D = 0{\cdot}25$. Thus the D-residue, the minor incorporated residue, is obliged to take a right-handed helical structure, following the major L-residue. This is the same conclusion as was first reached by Lundberg & Doty (1957) as a result of their kinetic study, and by Downie *et al.* (1957), in their study by optical rotation. Consequently, it can be concluded that the spiral regularity in the backbone chain of the α-helix is firmly retained, at least in the range of 0 to 0·25 fraction of D-residue.

Next let us consider the orientation of the polar groups in the D,L-copolymers, by taking the residual electric dipole moment μ_0 as a measure of the chain regularity, and by relating its plot against the L-fraction as shown in Fig. 5. A steady increase in μ_0 with increase in the D-fraction in this plot cannot be explained merely by the orientation of polar groups in the backbone chain because these already make their maximum contribution in the pure L-polymer, since their moment vectors are almost parallel with the helical axis (Wada, 1959*a*).

Consequently, the explanation can only be found in the orientation of the polar groups in the side-chains. In connection with this problem, Tsuboi (1961) has presented a model from a study of the polarized infrared absorption spectrum in which the C═O group in the side-chain is fixed at an angle of 53° to the helical axis rather than in a random orientation; this method cannot, however, give its direction. If this is the case, since C═O (dipole moment *ca.* 2·2 Debye units) is the main polar group of the side-chain, and can have a major effect on the molecular dipole moment, the direction will be such as to offset the electric dipole moment in the backbone-chain of the pure L-polymer. With such a model, the fact that μ_0 increases with the D-fraction can be explained as follows.

When a D-residue is incorporated in the right-handed α-helix, its side-chain will have a direction opposite to that of an L-residue in the same helix. Therefore, in the D,L-copolymer with a small amount of D-fraction ($f_D = 0$ to 0·25), the dipole moment in the side-chain of the D-residue will be in the same direction as that of its backbone-chain. In Fig. 8 the direction of the axial component of the electric dipole moment in the backbone and side-chains is shown.

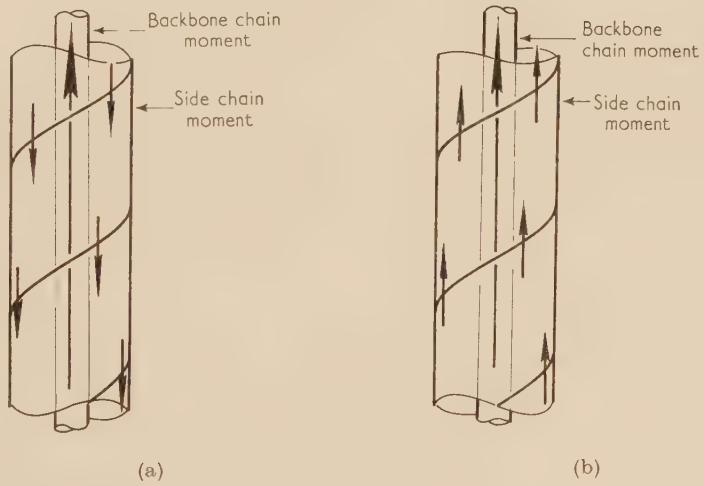


FIG. 8. An illustrative picture of the orientation of polar groups in the backbone and side-chain in the right-handed α-helix. (a) Helix of L-isomer. (b) Helix of D-isomer.

Unfortunately, the effect of the side-chain moment would not be offset completely in a meso-polymer even if a single screw-sense were present, because the direction of the side-chains of the L- and D-residue in a helix of given sense may not be symmetrical about the plane perpendicular to the helical axis, and because of the flexibility of the side-chain which originates in the internal rotation allowed in the carbon-carbon bond. By assuming that the side-chain is fairly rigid, however, the straight extrapolation to

$f_D = 0{\cdot}5$ (Fig. 5) can still be assumed to give an approximate value for the residual moment of the backbone-chain of the α-helix. When the effects of the moment induced by the reaction field (Wada, 1959*b*) of resonance and hydrogen bonding are taken into consideration, the moment 6·5 Debye thus obtained seems to be reasonable in such a medium of high dielectric constant ($\epsilon \fallingdotseq 10$, Wada, 1961).

We must next explain the behavior of μ_0 as it approaches $f_D = 0{\cdot}5$. It is obvious that the decrease in μ_0 above $f_D = 0{\cdot}35$ is attributable to some defect in the right-handed helix. However, the fairly large value of μ_0, which persists even in a polymer of f_D *ca.* 0·5, suggests that the unfolding of the ordered form is possible only at the end of the molecular chain, because the dipole moment is given by the vector sum of all contributions and a few breaks in the middle of the rigid backbone will produce a large decrease in μ_0. The existence of an unfolded portion at the end of the chain is also quite plausible because of the competition between L- and D-residues which exists until the helical configuration is produced after the initial stage of polymerization, as described above. If likewise there are disordered points in the molecule, they must either be few in number or rigid.

In any case, the experimental results suggest that a considerable number of regular configurations still exist even in a copolymer in which L- and D-residues are equally mixed.

Furthermore, in comparing the dielectric data with the results of the optical rotatory dispersion, it can be seen that the change in $(H_r - H_1)$ begins before the appearance of the defect in the regularity of the polar groups. This is another reason to believe that the material will be composed of right-handed helices of L-residue-rich polymer and left-handed helices of D-residue-rich polymer as would be expected from the previous kinetic considerations, although the formation of some kind of periodic configuration by a stoichiometric combination of L- and D-residues cannot be ruled out.

The discrepancy between the dielectric result and the optical rotatory dispersion mentioned above can be explained by both of these possibilities, but the first is much the more probable because no significant difference has been observed between the backbone structure of enantiomorphic and racemic polymers in the infrared dichroism (Tsuboi, Miyazawa & Wada, 1961) and X-ray diffraction study (Tsuboi & Wada, 1961). Details of these studies will be given in a later report.

One point to be noted is that the minimum in μ_0 does not occur at $f_D = 0{\cdot}5$ within the limit of experimental error, although the figure should be symmetrical about this axis. Both the starting materials (the L- and D-isomers) are pure (99·9%), but it is quite possible that this discrepancy is due to the asymmetry in the synthesizing process of L- and D-NCA, or else due to impurity, since even small contaminants may have a considerable effect on the electric dipole moment of a rigid molecule in such an unstable condition.

REFERENCES

Blout, E. R. & Karlson, R. H. (1956). *J. Amer. Chem. Soc.* **78**, 941.
Blout, E. R. & Miyazawa, T. (1961). In preparation.
Bradbury, E. M., Downie, A. R., Elliott, A. & Hanby, W. E. (1960). *Proc. Roy. Soc.* A, **259**, 110.
Doty, P. & Lundberg, R. D. (1956). *J. Amer. Chem. Soc.* **78**, 4810.
Doty, P. & Lundberg, R. D. (1957*a*). *Proc. Nat. Acad. Sci., Wash.* **43**, 213.

Doty, P. & Lundberg, R. D. (1957*b*). *J. Amer. Chem. Soc.* **79**, 2338.
Doty, P., Bradbury, J. H. & Holtzer, A. M. (1956). *J. Amer. Chem. Soc.* **78**, 947.
Downie, A. R., Elliott, A., Hanby, W. E. & Malcolm, B. R. (1957). *Proc. Roy. Soc.* A, **242**, 325.
Flory, P. J. (1953). *Principles of Polymer Chemistry.* New York: Cornell University Press.
Idelson, M. & Blout, E. R. (1958). *J. Amer. Chem. Soc.* **80**, 2387.
Imahori, K. (1960). *Biochim. biophys. Acta*, **37**, 336.
Lundberg, R. D. & Doty, P. (1957). *J. Amer. Chem. Soc.* **79**, 3961.
Miyazawa, T. & Blout, E. R. (1961). *J. Amer. Chem. Soc.* **83**, 712.
Moffitt, W. & Yang, J. T. (1956). *Proc. Nat. Acad. Sci., Wash.* **42**, 596.
Tsuboi, M. (1961). *J. Polymer Sci.* in the press.
Tsuboi, T. & Wada, A. (1961). In preparation.
Tsuboi, M., Miyazawa, T. & Wada, A. (1961). In preparation.
Wada, A. (1959*a*). *J. Chem. Phys.* **30**, 328.
Wada, A. (1959*b*). *J. Chem. Phys.* **31**, 495.
Wada, A. (1960*a*). *Bull. Chem. Soc. Japan*, **33**, 822.
Wada, A. (1960*b*). *J. Polymer Sci.* **45**, 145.
Wada, A. (1961). In preparation.
Wada, A., Tsuboi, M. & Konishi, E. (1961). *J. Phys. Chem.* in the press.
Yang, J. T. & Doty, P. (1957). *J. Amer. Chem. Soc.* **79**, 761.

J. Mol. Biol. (1961) **3**, 520–531

Thermodynamic Properties of Resilin, a Rubber-like Protein

Torkel Weis-Fogh

Zoophysiological Laboratory B, University of Copenhagen, Denmark †

and

Department of Zoology, Cambridge University, England

Previous experiments on the thermo-elastic behaviour of rubbery proteins are difficult to interpret for two reasons: the degree of swelling changes with temperature and the intricate structure of most samples renders them unsuited for accurate analyses. Resilin is an insoluble structural protein, recently discovered in arthropod cuticle. Pure samples of unstrained resilin are completely isotropic and are remarkable for their high reversible deformability, complete lack of flow and perfect elastic recovery even after prolonged strain. A small elastic tendon in dragonflies (Odonata) made it possible in a simple way to study its mechanical and thermo-elastic properties.

When swollen with water at pH 7, tendon resilin can be compressed almost to one third and extended to three times its unstrained length before breaking (at 30 kg/cm^2 unstrained, swollen area). The elastic modulus is 6 to 7 kg/cm^2, as in most rubbers. At constant length, the elastic force increases linearly with the absolute temperature over a considerable range of length and of temperature (20° to 80°C). A detailed analysis proved that the force originates from changes in entropy caused by the strain and only to a minor extent from changes in internal energy.

This is considered the first unambiguous proof for the rubber-like nature of any structural protein. It can be inferred that a piece of resilin is a three-dimensional network of highly flexible, cross-linked polypeptide chains in vivid thermal agitation. It has the character of a giant molecule with no inherent restrictions as to size or shape. Some biological implications are briefly discussed.

1. Introduction

There is much circumstantial evidence in favour of the belief that elastin and some other structural proteins exist in the rubber-like state of matter (see reviews by Wöhlisch, 1939; Astbury, 1940; Meyer, 1950; Kendrew, 1954), but it will be shown that the crucial evidence from thermodynamic experiments is ambiguous. The discovery of a new type of highly elastic structural protein, *resilin* (Weis-Fogh, 1960; Bailey & Weis-Fogh, 1961), made it possible to re-investigate the problem by using more suitable samples. Since the thermodynamic experiments reported here gave a clear-cut answer, it was legitimate to interpret in detail the mechanical and optical properties of resilin according to the kinetic theories of rubber elasticity and so to obtain valuable information about the molecular structure of at least one rubber-like protein (Weis-Fogh, in preparation).

Resilin is the essential elastic component in certain mechanical springs in the cuticle of arthropods (cf. the Latin *resilire*, to spring back) and it has a number of

† present address.

properties which make it ideally suited for this type of investigation (Weis-Fogh, 1959; 1960). When swollen with water, it exhibits typical long-range elasticity, it is heat-stable up to 140°C, cross-linked by very stable covalent bonds and insoluble in all solvents that do not break the peptide backbone but it is easily digested by all kinds of proteolytic enzymes. Moreover, it is secreted as thick continuous layers by the epidermis in a pure form unmixed with other proteins, lipids or polysaccharides and spatially separated from chitin, the only other structural component of the rubber-like cuticle (Weis-Fogh, 1960; Bailey & Weis-Fogh, 1961). It is thus possible to obtain native resilin in the form of relatively large, homogeneous samples, 0·1 mm thick and 1 mm long. Such samples are hyaline, devoid of colour, and mechanically as well as optically isotropic in all directions and at all degrees of swelling. They are also devoid of structure (light and electron microscopy), and no trace of crystallinity has been found (X-ray diffraction) even in samples that were stretched near to the breaking point and then slowly dried (Elliott, Huxley & Weis-Fogh, in preparation). The amino acid composition is unique and different from that of elastin, elastoidin, silk fibroin and collagen, and about one third of the side chains carry polar groups (Bailey & Weis-Fogh, 1961). The rubberiness of elastin is generally thought to be intimately connected with the scarcity of polar groups (5%; cf. Lloyd & Garrod, 1946).

This paper will deal with experiments that show that the elastic force of resilin is connected with entropy changes caused by deformation, as in true rubbers, rather than with changes in internal energy, as in most solids.

2. Theory and Earlier Results

In order to avoid ambiguity it is essential to study the basic thermodynamic relations on which the conclusions are based and, in particular, to specify the conditions for which they are valid.

(a) *Thermodynamics and the experiment*

For a reversible transformation of an isolated system we have $\mathrm{d}U = T\mathrm{d}S - \delta W$, where U is the internal energy, T the absolute temperature, S the entropy, and W the work done by the system on the surroundings. Let l represent length, P pressure, V volume, and f the stress in the direction of the strain $\mathrm{d}l$. In the absence of chemical reactions, for an isolated mechanical system $\delta W = P\mathrm{d}V - f\mathrm{d}l$, but in ordinary rubbers $|P\mathrm{d}V| \ll |f\mathrm{d}l|$ so that the total work performed can be written

$$\delta W = -f\mathrm{d}l \tag{1}$$

and, consequently, $f\mathrm{d}l = \mathrm{d}U - T\mathrm{d}S$. According to the definition of Helmholtz free energy $F = U - TS$, we then have for an infinitesimal change $\mathrm{d}F$:

$$\mathrm{d}F = \mathrm{d}U - T\mathrm{d}S - S\mathrm{d}T = f\mathrm{d}l - S\mathrm{d}T \tag{2}$$

During isothermal conditions (T = constant) the change in free energy therefore is $\mathrm{d}F = \mathrm{d}U - T\mathrm{d}S = f\mathrm{d}l$, so that

$$f = \left(\frac{\partial F}{\partial l}\right)_T = \left(\frac{\partial U}{\partial l}\right)_T - T\left(\frac{\partial S}{\partial l}\right)_T \tag{3}$$

Since the right-hand side of equation (2) is a complete differential in l and T, it follows that:

$$\left(\frac{\partial f}{\partial T}\right)_l = -\left(\frac{\partial S}{\partial l}\right)_T \tag{4}$$

When this equation is inserted in (3), we get the well-known thermodynamic equation of state first applied to rubber by Wiegand & Snyder (1934):

$$f = \left(\frac{\partial U}{\partial l}\right)_T + T\left(\frac{\partial f}{\partial T}\right)_l \tag{5}$$

The experimental significance of equation (5) is that the quantity $(\partial f/\partial T)_l$ can be measured directly for any given value of l as the slope of the experimental curve relating isometric force (l = constant) and absolute temperature. For a given temperature, $(\partial U/\partial l)_T$ then follows uniquely, but it is relevant to point out that, in principle, its value is estimated as the ordinate of the tangent's intercept for $T = 0$. For ordinary materials, therefore, the experimental procedure involves considerable extrapolation so that a high degree of accuracy and/or a large range of temperature is essential. When, as in most cases considered here, the thermodynamic system consists of a given sample of structural protein swollen with water, the most difficult requirements to comply with are that the system (a) must be in equilibrium at every reading, (b) must be closed so that there is no exchange of matter with the surroundings, for instance no change of swelling, and (c) must be devoid of chemical change caused by heating or stretching. Moreover the pressure and volume should be constant or nearly so during the transformations; otherwise the changes in work are not adequately described by equation (1) and even at equilibrium a third unknown term must be added to the right-hand side of equation (5). This implies that an increase of the isometric force with temperature need not indicate the presence of a true Joule effect and *vice versa* that a decrease does not demonstrate its absence.

(b) *Experiments with rubbers*

Long-range elasticity does not in itself classify a material as a rubber in the physical sense of this term. Swollen wool keratin, for instance, can be stretched reversibly to twice its length but behaves as an ordinary solid from the thermodynamic point of view (Bull, 1945). Apart from long-range elasticity there are two distinguishing features of the true rubber-like state, corresponding to the two terms in equation (5); the actual chemical composition is immaterial. The elastic force should be due mainly to entropy changes caused by the strain so that the isometric force increases linearly with T (proportional to T in ideal rubber); the internal-energy term $(\partial U/\partial l)_T$ should be small and nearly constant over a large range of temperature and length (zero in ideal rubber). Figure 1(a) illustrates a typical set of experimental results obtained with lightly vulcanized *Hevea* rubber, re-drawn from Anthony, Caston & Guth (1942). The isometric force increases nearly proportionally with the absolute temperature. The small decrease seen at the lowest extension ($\alpha = 1{\cdot}03$) is accounted for by general thermal expansion of the homogeneous, unswollen sample. The data interpreted according to equation (5) and Fig. 1(b) show that the force-extension curve is determined mainly by the changes in entropy with strain (broken curve). The subordinate effect of changes in internal energy becomes even more obvious after correction for thermal expansion (dotted curves).

(c) *Experiments with structural proteins*

The principal difficulties with structural proteins are that they are polyelectrolytes; when dry, they never behave like simple rubbers but are solid, and they exhibit long-range elasticity only when swollen with water or other polar solvents. It is also

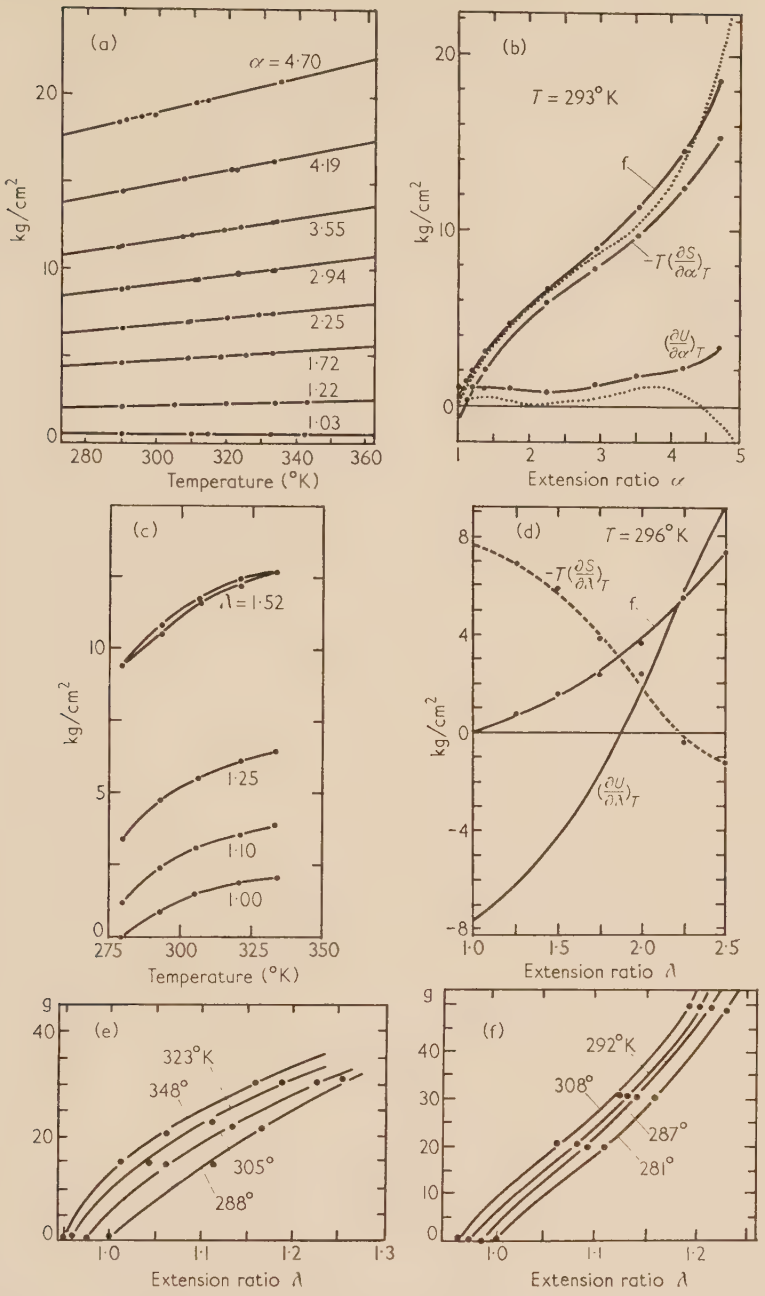


FIG. 1. Previous experimental results, re-drawn from the literature (see text). (a) isometric force of unswollen, vulcanized *Hevea* rubber against absolute temperature (Anthony *et al.*, 1942); (b) same data analysed according to equation (5), dotted curves after correction for thermal expansion; (c) isometric force of water-swollen ox *ligamentum nuchae*, i.e. mainly of elastin, against absolute temperature (Meyer & Ferri, 1936); (d) the force extension curve of ox *ligamentum nuchae* resolved into its apparent entropy term and internal-energy term (Wöhlisch, Weitnauer, Grüning & Rohrbach; 1943); (e) length of collagen-free aorta elastin in water extended by isotonic loads (ordinate) at different temperatures (Lloyd & Garrod; 1946); (f) same for silk fibroin swollen and super-contracted in formic acid (Lloyd & Garrod, 1946).

The definitions of the extension ratios α and λ are given on page 525.

difficult to obtain homogeneous samples at all suited for experiments. This does not apply to resilin. Since the affinity between sample and swelling medium may vary with temperature (which always seems to be the case with water) the system is neither closed nor purely mechanical. As in all swollen rubbers (cf. Treloar, 1958) a minor complication is caused by changes in hydrostatic pressure and thereby in swelling as a consequence of stretching.

In the case of elastin, the failure of earlier attempts to obtain convincing results seems to be due partly to the intricate microscopic structure and composition of available 'elastin' preparations (cf. Hall, Reed & Tunbridge, 1952; Cruise, 1957) and partly to neglect of swelling phenomena. In 1914, McCartney had shown that the total volume of samples from ox *ligamentum nuchae* in Ringer's solution decreases reversibly when heated both before and after removal of collagen (by means of pepsin) by approximately 20% from 15° to 70°C. Similar results were obtained by Wöhlisch (1932) who found that the shrinkage is not even isotropic. In accordance with this Meyer & Ferri (1936) determined that the isometric force of a suspended ox ligament increased with temperature not only in the strained sample but also in the unstrained one ($\lambda = 1{\cdot}0$), as is seen in the lower curve of Fig. 2(c). The latter increase was 2 kg/cm² which is an appreciable fraction of the change even at the highest extension for which they claim rubber-like behaviour ($\lambda = 1{\cdot}5$). They explained this "anomaly" as due to the presence of collagen but we have seen that this can only be part of the explanation. Moreover Fig. 1(e) (from Lloyd & Garrod, 1946) shows that the length of unstrained, collagen-free elastin from ox aorta varies considerably within the temperature range 315° to 335° K where, according to Meyer & Ferri, elastin should behave like rubber. However, the scarcity of experimental data and the changes in swollen volume prevent further analysis in terms of internal energy and entropy.

Such an attempt was made by Wöhlisch, Weitnauer, Grüning & Rohrbach (1943), also on samples from the elastic ligament of the ox. Although they claim that elastin is to be regarded as an imperfect rubber their results are, nevertheless, discouraging since the two deduced terms in equation (5) were found to vary in an almost opposite direction to the variation known from natural and synthetic rubbers (Fig. 1(d)). If these results were true one should be very hesitant to include swollen elastin among rubber-like materials. According to Lloyd & Garrod's experiments (1946) with collagen and with silk shrunk in formic acid (Fig. 1(f)) this also applies to other structural proteins in the so-called supercontracted state.

3. Materials and Methods

(a) *The elastic tendon*

Nature has provided an almost ideal sample of resilin. It consists of a cylindrical, highly extensible, swollen part which occupies the middle third of the otherwise tough tendon of the small pleuro-subalar muscles in dragonflies (Weis-Fogh, 1960). It is a hollow invagination from the body surface covered by a single layer of epidermal cells which can easily be stripped off. The chitinous ends are thin-walled, crystalline and almost inextensible while the stripped swollen part is thick-walled, isotropic and transparent like jelly and so elastic that it can be extended to three times the unstrained length λ_0 and still snap back to λ_0 without any permanent deformation (Fig. 2(a) & (b)). The swollen part, i.e. the elastic tendon proper, is shaped like a piece of rubber tubing and, in the metathorax of *Aeshna cyanea* Müll. and *A. grandis* L., it is a regular cylinder 0·8 mm long and 0·15 mm in diameter, pierced along the axis by the "central canal" having a circular bore of 0·03 mm

diameter. These dimensions refer to pH 7 but the elastic part swells isotropically in alkaline and shrinks in acid solutions with a minimum at pH 4. It contains about 6 μg of dry matter. The elastic tendon is covered by an extremely thin chitinous membrane (less than 0·2 μ), probably of a felt-like texture, while the bore is lined by an equally thin wavy membrane of epicuticle; neither membrane contributes to the elastic force by significant amounts. Within the accuracy of measurement both the mechanical and the optical effects are therefore due to the homogeneous, pure resilin which fills up the entire space between the two membranes and which is anchored to the tough parts by means of a short bell-shaped system of chitin lamellae at either end.

The elastic tendon goes solid and glass-like when dried but it re-swells and resumes all its former properties in less than a minute after being placed in water or buffer solution. This also happens after treatment in aqueous alcohols and after total dehydration in absolute methanol or ethanol in which the unswollen, unstrained length α_0 was measured. The extension ratio α is the strained length in units of α_0. The degree of swelling is given by the volume fraction v = (unswollen volume)/(swollen volume) and was estimated from microscopic measurements of total length and of inner and outer diameters (measured at several points along the fibre). The extension ratio λ, defined as the length in units of the unstrained, swollen length λ_0, is then $\lambda = \alpha v^{\frac{1}{3}}$.

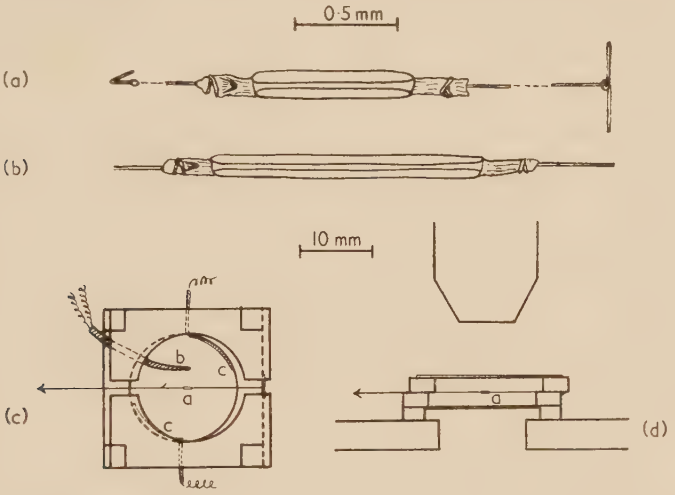


Fig. 2. (a) Unstrained *Aeshna* tendon mounted between nylon strings, pH 7; (b) same tendon stretched; (c) observation chamber with tendon, a, thermocouple, b, and electrodes, c for electric heating of water, as viewed from above and from the side (d).

(b) *Suspension, observation and temperature*

A single-stranded nylon fibre (0·02 mm diameter) was tied round each of the tough ends in such a way that the central canal was left open to allow free passage of liquid in and out (Fig. 2(a)). A small hook of platinum-iridium alloy (0·05 mm diameter) was fastened to one fibre and a transverse bar of the same material to the other. By pulling the nylon string the tendon was extended to twice its unstrained length (Fig. 2(b)), dehydrated through the alcohol series and dried in air. The taut knots were secured in the now centered position by means of tiny drops of celluloid in acetone and left to dry. On immersion in water, the re-swollen tendon remained centered at all degrees of extension, i.e. the tensile stress was equally distributed over the cross-sectional area, as judged by the stress birefringence. Control preparations were made in which ethanol and drying were avoided. The elastic properties were not influenced but the centering was less perfect. Small characteristic irregularities on the surface and/or minute pieces of platinum placed in the central canal served as reference points for length measurements.

The observation chamber (Fig. 2(c) & (d)) consists of two square glass plates with circular holes, glued together with Araldite and provided with small glass feet. The upper

plate has two slits for the nylon strings and the tendon is placed in water between two coverslips loosely stuck to the upper and lower surfaces by means of petroleum jelly. This prevents thermal stresses in the coverslips from influencing the estimates of birefringence. Two large platinum electrodes are placed opposite each other at the edge so that the chamber can be heated electrically and almost uniformly by means of an adjustable 3·3 kc/s generator. The temperature was read by means of a small thermocouple placed in the immediate neighbourhood of the tendon. Since the experiments took place in a constant-temperature room (18·0 $\pm$ 0·1°C), the temperature of the tendon could be adjusted to any desired value up to 90°C ($\pm$ 0·2°) in less than 5 min.

Under the polarizing microscope (E. Leitz) and with the first-order-red compensator, all linear dimensions could be read to within 2 μ by means of an eyepiece-screw micrometer. The cross-sectional area was calculated from measurements of the inner and outer diameters at several points along the tendon.

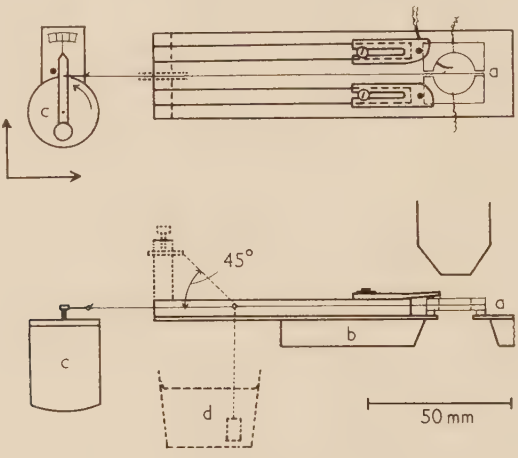


Fig. 3. Apparatus for stretching tendon either electro-magnetically (full lines) or by means of weights (broken lines).

a, Observation chamber;
b, Turntable for polarizing microscope;
c, Electromagnetic indicator mounted on micromanipulator;
d, Weight in dash pot.

(c) *Tensile force and extension*

The observation chamber was clamped above a hole cut in a plastic slide which fitted the turntable of the polarizing microscope (Fig. 3). The other end of the slide projected over the edge and carried a vertical support and a dash pot below (broken lines). The tendon was extended either by an adjustable horizontal pull from the hand of an electromagnetic indicator (Smith's Desynn Indicator, 24 V) or by means of weights, as indicated in Fig. 3. The latter procedure offers constant loading (isotonic) for an unlimited period of time and permits any orientation of the tendon relative to the planes of polarization. The electro-magnetic device was mounted on a Prior micromanipulator with adjustment facilities in three directions. It was used as a zero instrument so that the applied force is proportional to the d.c. current read on a calibrated micro-ammeter. In order to test the sample under dynamic conditions, a sinusoidal a.c. component (50 c/s) of variable amplitude was sometimes superimposed for short periods of time. The force was transmitted from the tendon by means of a single thick nylon fibre (0·1 mm diameter) and further by means of Constantan wire (0·12 mm diameter). The maximum force was 4·5 g and the accuracy of reading $\pm$ 0·005 g.

The apparatus is of the isotonic type and could be used for thermo-elastic experiments at constant length only because of the speed with which equilibrium is attained in the

small tendon. In fact, true isometric experiments are not possible because of the compliance of the tough parts at either end. The procedure was to select a certain extension and to keep it constant to within 0·5% of the total length by continuous adjustments of the current when the temperature changed, equilibrium being obtained within the period of manual adjustment, i.e. in less than 10 to 20 sec. Since equilibrium was attained so quickly at all strains and temperatures, it was not necessary to expose the tendon to any initial regime in order to obtain reversible and reproducible results. For most rubbers such a regime is essential.

4. Results

(a) *General properties*

In neutral buffer the volume fraction v is about 0·45 and the tensile strength is about 30 kg/cm² of the unstrained cross-sectional area (70 kg/cm² based on the dry area). The broken surfaces are nearly perpendicular to the axis and have a conchoidal texture. After breaking they fit exactly together and the re-assembled tendon keeps its former dimensions, indicating complete elastic recovery and absence of flow. In terms of the unstrained, swollen length λ_0, the tendon could be compressed at least down to 0·4 λ_0 and extended to at least $3\lambda_0$.

Elastic recovery and flow was also studied during prolonged isotonic loading. When a tendon was extended to about $\lambda = 2$ by means of a weight, the length remained constant (within 0·5%) for days and even weeks at 4°C; a crystal of thymol was added to the buffer to prevent bacterial growth. Upon unloading, the tendon immediately snapped back to its dimensions before the experiment and gave the same length-tension diagram as when first tested. When the electromagnet was switched on and off, it was found that the adjustment to any new length lasted less than 1 second so that no hysteresis could be observed. The rapid adjustment and the complete absence of thixotropy became apparent also when a sinusoidal strain (50 c/s, amplitude about 0·05 λ) was superimposed upon the static strain. The average extension and the average strain birefringence (estimated by means of a Berek compensator) both remained unchanged.

The load-extension diagram of a tendon from *Aeshna grandis* is shown in Fig. 4, measured first at 18°C (291°K) and then at 81°C (354°K). This tendon was not placed in phosphate buffer but in a neutral 1 : 1 mixture of distilled water and Cambridge tap water which had a suitable conductivity for the electric heating. In contrast to collagen and elastin, the dimensions of resilin in unbuffered water *increase* isotropically with increasing temperature, in this case by 4% in length and by 12% in volume. At constant extension, the elastic force should therefore decrease with increasing temperature if swelling phenomena were of major importance. However, the slope of the curve B is larger than that of curve A for *all* extensions up to $\alpha = 2{\cdot}5$; the two curves intersect at $\alpha = 1{\cdot}6$, the thermo-elastic point of inversion. When the same tendon was placed in 0·003 M-phosphate buffer at pH 6·7, heating from 18° to 80°C affected swelling in the same direction as in unbuffered water but it became reduced to only 1% by length and 3% by volume. The thermo-elastic point of inversion was $\alpha = 1{\cdot}4$.

It is clear that the elastic force increases with increasing temperature and that this property is not due to swelling phenomena. Moreover, the shape of the force-extension curves, the elastic modulus (6 to 7 kg/cm²), and the tensile strength are typical of true rubbers.

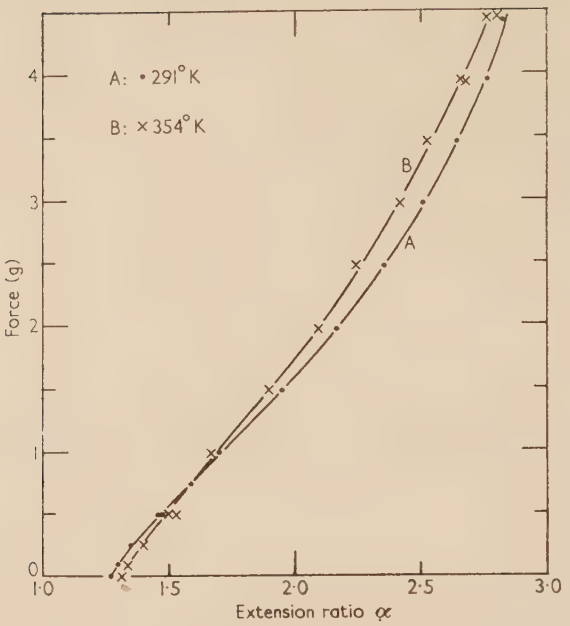


FIG. 4. Elastic tendon of *Aeshna grandis*: tensile force (ordinate) plotted against the extension ratio α (abscissa), measured at two temperatures. Unbuffered, neutral water.

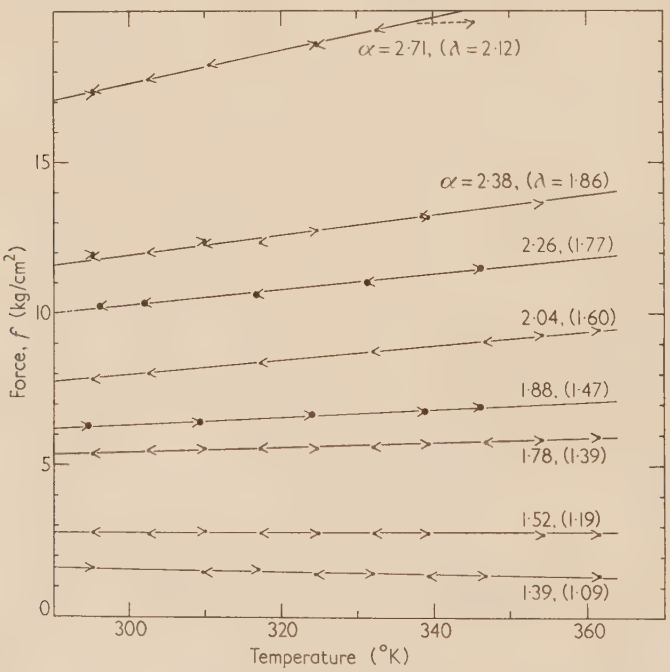


FIG. 5. Tensile force per cm² unstrained, swollen area (f, ordinate) at various constant lengths as a function of the absolute temperature (T, abscissa) in an elastic tendon from *Aeshna grandis*. The extension ratios are written against the curves, as based upon the unstrained, dry length (α) and on the unstrained, swollen length at pH 7 (λ, in brackets). Small symbols unbuffered water, large symbols water buffered at pH 7. The arrow heads indicate whether equilibrium was obtained during a sequence of rising or decreasing temperature.

(b) *Thermo-elastic properties*

Figure 5 demonstrates that the elastic force at constant length (ordinate) increases linearly with the absolute temperature (abscissa) over a considerable range of temperature and extension. The force is given in kg/cm² of the unstrained, swollen area at 291°K. The temperature was altered every 5 minutes either up or down, as indicated by the arrow heads. The changes were fully reversible within the accuracy of measurement. The straight lines, $f = a + bT$, were calculated according to the method of least squares and the standard deviations of the constants found. The experiments in Fig. 5 were performed on a tendon from *Aeshna grandis* placed either in unbuffered water (small points) or in phosphate buffer (large points), but it was not possible to measure any significant difference at extensions equal to or exceeding $\alpha = 1{\cdot}8$. Similar results were obtained with tendons from *Aeshna cyanea*.

The curves represent equilibrium situations in a reversible transformation; there was no indication of flow or hysteresis, and the effect of temperature on the swelling was but small. Moreover, experiments to be reported elsewhere (Weis-Fogh, in preparation) demonstrate that the elastic modulus as referred to the unswollen area remains practically unchanged and independent of the actual degree of swelling over the very large pH range from pH 1·8 to pH 12·3, indicating that no essential changes occur in the internal structure of the molecular chains, even under these drastic conditions. It is therefore legitimate to apply equation (5) to this protein. Consequently, the constants a and b were estimated for each extension directly from the regression lines. Since the force-temperature curves are linear and, with the units used in Fig. 5, a equals $(\partial U/\partial\lambda)_T$ and b equals $-(\partial S/\partial\lambda)_T$ for one cubic centimetre of swollen resilin.

It is highly significant that a turned out to be almost constant at all extensions, about 2 to 3 kg/cm². An analysis of the data from Fig. 5 is given in Fig. 6 in which the standard deviations are represented by vertical bars, the temperature being 300°K. The figure shows that the internal energy term is positive but small compared with the entropy term. In fact, it is the change in entropy with length that determines the shape of the force-extension curve and is responsible for the major part of the elastic force, more than 90% at $\lambda = 2{\cdot}0$ ($\alpha = 2{\cdot}6$). The two curves take the same general course as in vulcanized *Hevea* rubber (Fig. 1(b)) and if it had been possible to correct for thermal expansion, as in the case of rubber (dotted curves in Fig. 1(b)), the entropy term would have been greater and the internal energy term correspondingly smaller than appears from Fig. 6. It is of course not surprising that the absolute values are of the same order of magnitude in natural rubber and resilin, the elastic moduli being nearly the same.

TABLE 1

The constant a *in tendons from two dragonfly species. The tensile force ranged from* 1·4 *to* 20 *kg/cm² and the temperature from 292° to 352°K.*

		$\lambda = 1{\cdot}1$	1·2	1·4	1·6	1·9	2·6 (relative units)
A. grandis tendon:	a	2·6	2·9	3·0	0·8	2·5	— (kg/cm²)
A. cyanea tendon:	a	2·8	2·4	2·3	2·1	2·3	2·0 (kg/cm²)

The lower curve of Fig. 6 shows a minimum at $\lambda = 1{\cdot}6$ ($\alpha = 2{\cdot}1$) but, although statistically significant, it is probably due to some error or to minor changes in swelling (hydrostatic effects). It was absent in a tendon from *Aeshna cyanea*, as is seen in Table 1. Even when stretched to $\lambda = 2{\cdot}6$ ($\alpha = 3{\cdot}4$), a remained as small as 2 kg/cm^2, i.e. 10% of the elastic force at this length.

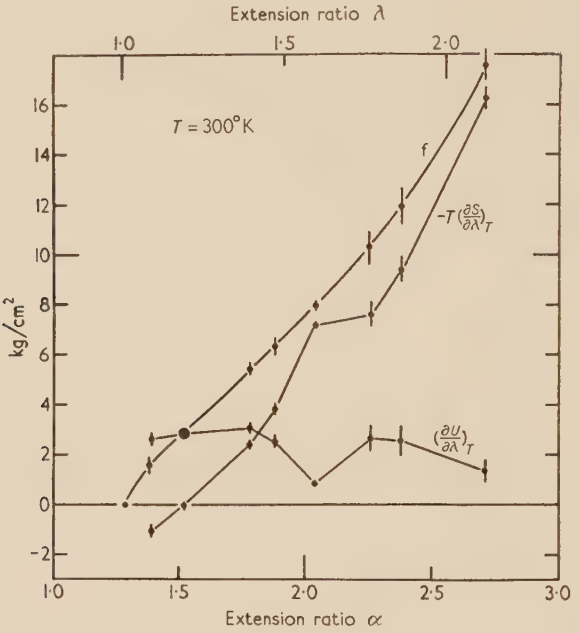


Fig. 6. The force-extension curve of the tendon in Fig. 5 resolved into the two components in equation (5), the entropy term and the internal energy term. The values refer to 1 cm^3 of resilin swollen with water at pH 7, the temperature being 300°K. The standard deviations are given by vertical bars.

5. Discussion

The kinetic interpretation of rubber elasticity offered by Meyer, von Susich & Valko in 1932 has stood the test of later experiments and forms the basis of most quantitative theories concerned with the mechanical and optical properties of rubber, with the effect of swelling, of temperature, crystallisation, etc. (see Treloar, 1958). Although they differ in approach and detail, the various kinetic theories rest on the basic assumption that a typical rubber consists of an isotropic network in three dimensions of long flexible chain molecules in vivid thermal agitation. The chains are kept fixed in the network by means of a few stable cross-linkages but otherwise they are free of each other. Since the individual links of a chain can rotate relative to their neighbours (flexibility) each chain as well as the entire assembly tends to take up the statistically most probable configuration. A deformation therefore leads to a change in entropy. At ordinary strains the tensile load on a chain is so small that a deformation hardly alters its energy content (internal energy), the main effect being a change in its configurational entropy and thereby in the entropy of the network.

We have seen that most preparations of elastin and other structural proteins which exhibit long-range elasticity do not behave like rubber but deviate so much from the simple pattern that the ordinary kinetic theories may not be applicable, at least not

for quantitative estimates. This does not mean that the proteins in question may not exist in the rubber-like state of matter because the deviations may be apparent rather than real. They may be caused by the intricate microscopic texture of the samples, by swelling phenomena, crystalline "impurities" etc. It is therefore important to have found that resilin which resembles elastin in so many respects (Weis-Fogh, 1960) can be regarded as a true rubber. Further information can then be obtained from simple physical experiments (Weis-Fogh, in preparation). It is of particular interest that the approximate number of cross-linking groups should be estimated since it facilitates research on their chemical nature. Both in elastin and in resilin they are of a covalent and extremely stable type and seem to differ from any known type (Weis-Fogh, 1960). It was also found that the birefringence depends only on the stress, as in true rubbers, and not on the degree of swelling so that it may be possible in a simple way to study many photo-elastic effects in microscopic structures like myofibrils, cilia and stroma proteins.

In some respects resilin appears to be superior to other rubbers, e.g. in its perfect elastic recovery and the complete lack of flow. These and other aspects will be discussed elsewhere. Here it is sufficient to recall that living organisms can produce perfect and highly deformable networks composed not of a single type of repeating unit but of 15 different amino acids (Bailey & Weis-Fogh, 1961) joined together into a three-dimensional fabric which has the character of a deformable giant molecule of unlimited size. Do similar structures play a role in the life of ordinary cells or during embryonic development?

I wish to thank Dr. O. Sten-Knudsen for many valuable suggestions, Sir James Gray, F.R.S., and Professor V. B. Wigglesworth, F.R.S., for their active interest. The experimental work was supported by the Department of Scientific and Industrial Research.

REFERENCES

Anthony, R. L., Caston, R. H. & Guth, E. (1942). *J. Phys. Chem.* **46**, 826.
Astbury, W. T. (1940). *J. Int. Soc. Leath. Chem.* **24**, 69.
Bailey, K. & Weis-Fogh, T. (1961). *Biochim. biophys. Acta*, **48**, 452.
Bull, H. B. (1945). *J. Amer. Chem. Soc.* **67**, 533.
Cruise, A. J. (1957). *Nature*, **179**, 674.
Hall, D. A., Reed, R. & Tunbridge, R. E. (1952). *Nature*, **170**, 264.
Kendrew, J. C. (1954). In *The Proteins*, ed. by H. Neurath & K. Bailey, II B, p. 845. New York: Academic Press.
Lloyd, D. J. & Garrod, M. (1946). *Fibrous Proteins, Soc. Dyers & Colourists Symp.* p. 24.
McCartney, J. E. M. (1914). *Quart. J. Exp. Physiol.* **7**, 103.
Meyer, K. H. (1950). *Natural and Synthetic High Polymers*. 2nd edition. New York & London: Interscience Publishers.
Meyer, K. H. & Ferri, C. (1936). *Pflüg. Arch. ges. Physiol.* **238**, 78.
Meyer, K. H., von Susich, G. & Valko, E. (1932). *Kolloidzschr.* **59**, 208.
Treloar, L. R. G. (1958). *The Physics of Rubber Elasticity*. 2nd edition. Oxford: The Clarendon Press.
Weis-Fogh, T. (1959). *Proc. XVth Internat. Congr. Zool.* 1958, p. 393.
Weis-Fogh, T. (1960). *J. Exp. Biol.* **37**, 889.
Wiegand, W. B. & Snyder, J. W. (1934). *I.R.I. Trans.* **10**, 234.
Wöhlisch, E. (1932). *Ergebn. Physiol.* **34**, 406.
Wöhlisch, E. (1939). *Kolloidzschr.* **89**, 239.
Wöhlisch, E., Weitnauer, H., Grüning, W. & Rohrbach, R. (1943). *Kolloidzschr.* **104**, 14.

J. Mol. Biol. (1961) **3**, 532–540

Coincidence and Protein Structure

J. Williams, J. B. Clegg

Department of Biochemistry, University of Cambridge, England

AND

Margaret O. Mutch

Mathematical Laboratory, University of Cambridge, England

(*Received 10 March 1961*)

The possibility that some of the regularities shown by amino acid sequences might be due to chance has been studied. Random sequences corresponding to ribonuclease, tobacco mosaic virus protein, trypsinogen, chymotrypsinogen and the A and B chains of haemoglobin were examined for the presence of intramolecular repeats of peptides and for peptide units present in more than one sequence. None of these proteins differed from random sequences with respect to intramolecular repeats. In the parts of their sequences now known the pairs of proteins trypsinogen and chymotrypsinogen, A chain and B chain of haemoglobin have more common peptide units than do the corresponding random sequences. The dipeptide frequencies in 13 proteins and peptides show no departure from randomness.

1. Introduction

In general, the naturally occurring amino acid sequences now known show no simple regularities. Yčas (1958) found no evidence from dipeptide frequencies that some amino acids are linked together more often than would be expected from chance. Similarly the amino acid compositions of proteins do not show correlations between pairs of amino acids (Augenstine, Branson & Carver, 1953). Since in these studies data from many proteins and peptides were pooled the results did not exclude the possibility that individual proteins might show preferred linkages. Šorm, in particular, claims that proteins are non-random in that (a) some peptides occur in more than one protein, (b) some peptides occur more than once in a single protein and (c) some peptides occur with different frequencies from those expected from chance (Šorm *et al.*, 1957; Šorm, 1959).

Šorm (1954; Šorm *et al.*, 1957) suggested that even where complete sequences are not known the number of peptides common to two proteins can be used to show similarity of their primary structures and he has applied the method to the comparison of trypsinogen and chymotrypsinogen. Thus, Keil *et al.* (1959) identified about half of the peptide bonds in the two zymogens and their data showed that 2 tetrapeptides, 7 tripeptides and 46 dipeptides occur in both proteins and that the dipeptide Ser.Gly occurs at least four times in each protein. They concluded that these proteins, which are related in function and site of origin, also have far-reaching structural similarities.

On the other hand, peptides may also be common to widely different proteins. Thus, the tetrapeptide Ala.Lys.Phe.Glu occurs in lysozyme and ribonuclease. Anfinsen (1959) points out that this might be due to "recurring chemical fine structure in the genetic material of the nucleus. Alternatively, various common peptide intermediates that have structures suited to specific functional requirements might be involved in the assembly of different classes of proteins."

Silk fibroin (Lucas, Shaw & Smith, 1958) and protamines (Monier & Jutisz, 1954) probably contain repeated subunits but, according to Šorm, regularities are also present in other proteins. In ribonuclease some dipeptides and tripeptides occur more than once and the distribution of lysine residues is said to show a periodicity which is possibly related to the turns of the α-helix (Šorm & Keil, 1958). These authors conclude that "there are regularities even in relatively small polypeptide units which cannot be attributed to coincidence." Nevertheless, they made no attempt to study chance regularities in linear sequences.

Šorm (1959) has tabulated 659 dipeptides from a collection of unspecified proteins and has noted that in some cases peptide AB occurred more often than the reverse peptide BA. According to Šorm these differences cannot be accidental (zufällige) because, from probability theory, AB should equal BA. This argument would only be valid for very large samples and Šorm gives no estimate of the amount of variation which might be due to random sampling.

Since Šorm appears to have dismissed without study the possibility that the regularities of protein structure which he has noted might be merely random we have tried to obtain some information on this point. No convenient theoretical treatment seems to be available so we have generated and analysed random sequences of appropriate sizes and compositions, using the electronic computer (EDSAC 2) in the Mathematical Laboratory, Cambridge.

2. Sequences Common to Two Proteins

Six infinitely large pools of symbols were used initially. Their compositions and the lengths of the sequences drawn from them were the same as the compositions and lengths of the following naturally occurring sequences:

(i) ribonuclease—R (Hirs, Moore & Stein, 1960);
(ii) tobacco mosaic virus protein—TMV (Anderer, Uhlig, Weber & Schramm, 1960);
(iii) a partial sequence accounting for 127 of the bonds in chymotrypsinogen—C (Keil *et al.*, 1959);
(iv) a partial sequence accounting for 96 of the bonds in trypsinogen—T (Keil *et al.*, 1959);
(v) a partial sequence accounting for 57 of the bonds in the A chain of human haemoglobin—HbA (Braunitzer, Rudloff, Hilse, Liebold & Müller, 1960);
(vi) a partial sequence accounting for 101 of the bonds in the B chain of human haemoglobin—HbB (Braunitzer, Hilschmann, Hilse, Liebold & Müller, 1960).

In addition two extra pools were used with the compositions of chymotrypsinogen—C_a (Wilcox, Cohen & Tan, 1957) and trypsinogen—T_a (Šorm, 1958). The lengths of the sequences drawn from these pools were 243 and 227 respectively.

A set of fifty random sequences was drawn from each pool and the following sets were compared:

(i) R—TMV;
(ii) R—C;
(iii) R—T;
(iv) TMV—C;
(v) TMV—T;
(vi) T —C;
(vii) T_a—C_a;
(viii) HbA—HbB;
(ix) R—HbB.

In the haemoglobin sequences the members of the pairs of residues glutamic acid-glutamine, aspartic acid-asparagine, leucine-*iso*leucine were not differentiated and for comparison (ix) the ribonuclease pool was modified so as to be the same as the HbB pool in this respect.

In making the comparisons each member of one set of random sequences was paired with one member of the other set. Each sequence was reduced to overlapping lengths of 2, 3 and 4 symbols to represent di-, tri- and tetrapeptides. Thus, the sequence ABCDE gives AB BC CD DE, ABC BCD CDE and ABCD BCDE. Peptides which were present in both members of a pair were counted and the mean values ($\bar{x}$) and the standard deviations (S.D.) for each set of comparisons are shown in Table 1. The figures in brackets in Table 1 are the numbers of common peptides observed in the naturally occurring sequences.

TABLE 1

Comparison	Common dipeptides $\bar{x}$	S.D.	Common tripeptides $\bar{x}$	S.D.	Common tetrapeptides $\bar{x}$	S.D.
T—C	27·5	3·4	3·1	1·9	0·2	0·5
	(46) †		(7) ‡		(2) †	
T_a—C_a	75·5	4·7	15·2	3·5	1·5	1·3
	(46) †		(7) ‡		(2)	
HbA—HbB	20·8	3·5	2·6	1·6	0·2	0·5
	(27)		(8) †		(2) †	
TMV—C	35·7	3·9	4·4	2·2	0·4	0·7
	(44) ‡		(3)		(1)	
TMV—T	26·6	4·1	3·4	2·3	0·1	0·4
	(23)		(2)		(0)	
R—HbB	27·6	4·6	3·6	2·2	0·3	0·5
	(33)		(3)		(0)	
R—TMV	29·7	4·7	3·3	2·1	0·1	0·4
	(31)		(4)		(0)	
R—C	30·0	4·1	2·9	2·2	0·2	0·4
	(27)		(2)		(0)	
R—T	23·1	3·6	3·1	1·8	0·2	0·5
	(18)		(2)		(0)	

The numbers of peptides common to random sequences drawn from eight pools. Mean values ($\bar{x}$) and standard deviations (S.D.) for fifty pairs. Figures for the corresponding naturally occurring sequences are given in brackets.

† $P < 0{\cdot}001$. ‡ $P < 0{\cdot}05$.

The known portions of the sequences of trypsinogen and chymotrypsinogen have more common dipeptides ($P < 0{\cdot}001$), tripeptides ($P = 0{\cdot}04$) and tetrapeptides ($P < 0{\cdot}001$) than do the corresponding random sequences T and C. The random sequences which correspond to the complete zymogen molecules (T_a and C_a), however, have as many common tetrapeptides as do the parts of the sequences so far known and more common di- and tripeptides ($P < 0{\cdot}001$). These data suggest that the zymogens possess real (i.e. non-random) structural similarity, but it is not certain that the complete sequences will have more common peptides than would be expected from chance alone.

The known parts of the A and B chains of human haemoglobin have more common tripeptides ($P < 0{\cdot}001$) and tetrapeptides ($P < 0{\cdot}001$) than do the corresponding

random sequences but the number of common dipeptides is not significantly different from the mean for the random sequences. According to Braunitzer, Liebold, Müller & Rudloff (1960) homology of the A and B chains of haemoglobin is shown by the occurrence of some residues in matching positions in the two chains. Clearly, homology of this sort need not be reflected in the numbers of peptides common to the two chains, as shown in this case by the number of common dipeptides.

With the exception of the number of dipeptides common to TMV protein and chymotrypsinogen (P 0·03 to 0·04) the numbers of common peptides found in the other comparisons do not differ significantly from the means for the corresponding random sequences.

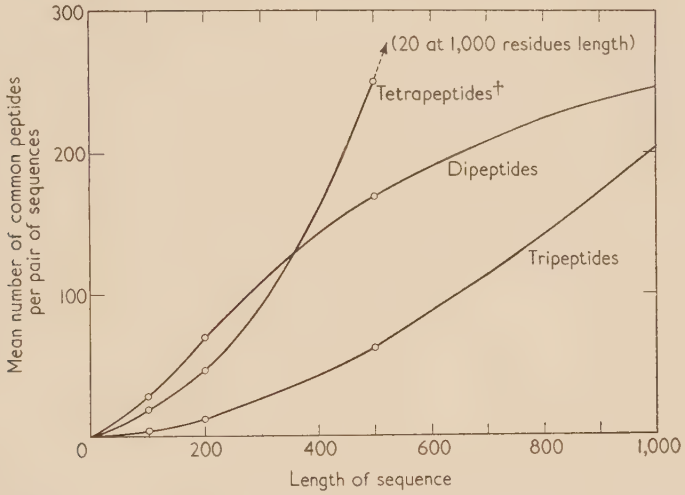


Fig 1. Mean numbers of peptides common to pairs of sequences of lengths 100, 200, 500 and 1000 drawn from the pool with average composition.

† For the values relating to tetrapeptides divide the ordinate scale by 50.

Since most proteins are bigger than the six proteins considered so far we tried to find the effect of chain length on the chance occurrence of common peptides between pairs of sequences. A pool of symbols with an average composition (from the compositions of 23 non-structural proteins given by Yčas, 1958) was formed and 50 pairs of random sequences containing 100, 200, 500 and 1,000 symbols were produced. The common peptides found for each length are shown in Table 2 and graphically in Fig. 1.

Table 2

Length of sequence	Common dipeptides		Common tripeptides		Common tetrapeptides	
	$\bar{x}$	S.D.	$\bar{x}$	S.D.	$\bar{x}$	S.D.
100	26·5	4·1	2·6	1·9	0·4	0·7
200	69·2	5·3	10·6	2·9	0·9	0·9
500	168·0	7·8	61·7	6·3	4·9	2·3
1000	245·8	5·5	203·7	12·1	19·9	5·3

The numbers of peptides common to pairs of random sequences drawn from a pool of average composition. Mean values and standard deviations for fifty pairs.

Over this range of lengths the numbers of common tripeptides and tetrapeptides occurring randomly rise steeply with increasing chain length. From chance alone, large single-chain proteins with molecular weights from 50,000 to 150,000 must be expected to have many short sequences in common. For this reason, it is not by itself meaningful that some tripeptides or tetrapeptides have been found in more than one protein (e.g. Šorm, 1959; Karadzova, Keil & Šorm, 1960).

3. Sequences Repeated within a Protein

Fifty random sequences were drawn from each of the six pools (R, TMV, T, C, HbA, HbB) and the number of times each di-, tri- and tetrapeptide occurred in each sequence was noted. The means ($\bar{x}$) and standard deviations (S.D.) for each set of fifty sequences are shown in Table 3. The numbers of repeated peptides found in the naturally occurring sequences are shown in brackets.

TABLE 3

	Occurrences	Dipeptides $\bar{x}$	S.D.		Tripeptides $\bar{x}$	S.D.		Tetrapeptides $\bar{x}$	S.D.	
R	2 times	17·7	3·7	(13)	2·0	1·7	(2)	0·2	0·7	(0)
	3 ,,	3·8	1·9	(4)	0·0	0·1	(0)	—	—	—
	4 ,,	0·6	0·8	(0)	—	—	—	—	—	—
TMV	2 ,,	24·0	3·1	(24)	3·9	1·9	(4)	0·3	0·4	(0)
	3 ,,	6·9	2·5	(8)	0·0	0·2	(0)	—	—	—
	4 ,,	2·2	1·4	(2)	—	—	—	—	—	—
	5 ,,	0·4	0·7	(0)	—	—	—	—	—	—
C	2 ,,	19·2	4·1	(15)	3·1	1·8	(0)	0·1	0·3	(0)
	3 ,,	4·4	1·8	(2)	0·1	0·3	(0)	—	—	—
	4 ,,	0·9	1·0	(2)	—	—	—	—	—	—
	5 ,,	0·2	0·4	(0)	—	—	—	—	—	—
T	2 ,,	12·6	3·5	(7)	2·2	1·9	(0)	0·2	0·4	(0)
	3 ,,	3·4	1·7	(2)	0·1	0·2	(0)	—	—	—
	4 ,,	1·0	0·9	(0)	—	—	—	—	—	—
	5 ,,	0·2	0·4	(1)	—	—	—	—	—	—
HbA	2 ,,	6·3	1·7	(7)	0·8	1·0	(0)	0·1	0·3	(0)
	3 ,,	1·3	1·0	(0)	—	—	—	—	—	—
HbB	2 ,,	15·6	3·2	(13)	2·7	1·5	(1)	0·3	0·7	(0)
	3 ,,	4·5	1·9	(3)	0·02	0·1	(0)	—	—	—

Repetitions of di-, tri- and tetrapeptides in random sequences corresponding to ribonuclease, TMV protein, chymotrypsinogen, trypsinogen, A and B chains of human haemoglobin. Means and standard deviations for 50 sequences of each type. Figures in brackets are the values for the naturally occurring sequences.

The numbers of repetitions observed do not differ significantly from the means for the corresponding random sequences and in this respect, therefore, these six sequences are random.

Table 4 shows the numbers of repeated peptides found in random sequences of lengths 100, 200, 500 and 1,000 drawn from the average pool. For each length 100 samples were used. In chains containing upwards of 500 residues many tripeptides and some tetrapeptides occur more than once.

TABLE 4

Occurrences	Length of Sequences: 100 $\bar{x}$	100 S.D.	200 $\bar{x}$	200 S.D.	500 $\bar{x}$	500 S.D.	1000 $\bar{x}$	1000 S.D.
Repetitions of Dipeptides								
2 times	12·5	3·1	31·7	4·9	59·6	7·3	57·8	6·5
3 ,,	2·3	1·2	10·0	2·9	33·9	4·6	44·7	6·6
4 ,,	0·3	0·5	2·4	1·4	19·1	3·7	34·0	5·5
5 ,,	0·1	0·3	0·6	0·9	10·0	2·9	24·2	5·4
6 ,,	—	—	0·2	0·4	4·9	2·1	17·5	3·7
7 ,,	—	—	—	—	2·0	1·4	12·6	3·3
8 ,,	—	—	—	—	1·0	1·1	8·3	2·6
9 ,,	—	—	—	—	0·4	0·6	6·0	2·1
10 ,,	—	—	—	—	0·2	0·5	4·2	1·9
Repetitions of Tripeptides								
2 ,,	1·6	1·1	5·3	2·3	31·7	5·5	108·3	9·9
3 ,,	0·1	0·3	0·2	0·5	1·3	1·1	9·6	3·7
4 ,,	—	—	—	—	0·3	0·3	0·7	0·8
5 ,,	—	—	—	—	—	—	0·2	0·4
Repetitions of Tetrapeptides								
2 ,,	0·1	0·3	0·2	0·5	2·5	1·7	10·4	4·1
3 ,,	—	—	—	—	—	—	0·1	0·3

Repetitions of di-, tri- and tetrapeptides in random sequences of lengths 100, 200, 500 and 1000 drawn from the pool with average composition. Means and standard deviations for 50 sequences of each length.

4. Dipeptide Frequencies

Šorm (1959) collected 659 dipeptides from unspecified proteins and noted that although in some cases the categories AB and BA were equally represented in other cases they were not. Thus:

	AB	BA		AB	BA
Val.Glu	10	1	Ileu.Val	4	0
Ser.Arg	12	1	Thr.Pro	4	0
Cys.Ala	7	1	Pro.Val	6	0

According to Šorm these differences could hardly be accidental, and Pasynskiĭ (1959) has suggested a thermodynamic explanation for them.

To illustrate the possibility of random differences we drew four lots of 500 dipeptides from the average pool. The biggest differences are shown in the following table. They suggest that some, at least, of the differences noted by Šorm might be accidental.

	AB	BA		AB	BA		AB	BA		AB	BA
Glu.Lys	7	2	Ser.Asp	8	1	Lys.Asp	10	3	Lys.Ala	6	1
Lys.Asp	7	2	Glu.Ileu	6	0	Asp.Glu	9	4	Ala.Ser	6	2
Val.Glu	5	1	Asp.Thr	6	1	Ser.Leu	5	0	Glu.Asp	8	4
						Glu.Tyr	6	2			
						Asp.Leu	8	4			

In Table 5 we list the dipeptide frequencies in the following proteins and peptides :

(i) insulin (Ryle, Sanger, Smith & Kitai, 1955)
(ii) ribonuclease (Hirs *et al.*, 1960)
(iii) lysozyme (Jollés & Jollés, 1960)
(iv) glucagon (Bromer, Sinn, Staub & Behrens, 1956)
(v) papain (Light & Smith, 1960)
(vi) peptides A and B from fibrinogen (Folk & Gladner, 1960; Folk, Gladner & Levin, 1959)
(vii) hypertensin (Skeggs, Lentz, Kahn, Shumway & Woods, 1956)
(viii) β-corticotropin (Bell, 1954)
(ix) TMV protein (Anderer *et al.*, 1960)
(x) human haemoglobin B chain (Braunitzer, Hilschmann, Hilse, Liebold & Müller, 1960)
(xi) chymotrypsinogen (Keil *et al.*, 1959)
(xii) trypsinogen (Keil *et al.*, 1959)
(xiii) peptide from beef cytochrome *c* (Tuppy, 1958).

TABLE 5

First residue	Second residue: ser	ala	val	gly	leu	asp	glu	thr	lys	arg	pro	phe	cys	tyr	asn	gln	ileu	his	met	try	
ser	9	7	2	12	5	6	1	5	4	5	1	5	3	5	4	5	2	1	1		83
ala	9	13	4	5	7	6	2	4	4	3		3	3	1	2	3	2	5	1	1	78
val	5	5	3	6	3	8	7	5	3	4	3	3	8	3	1	5		3		2	77
gly	6	5	7	6	6	6	6	4	7	3	5	1	3	3	4		3				75
leu	6	6	6	5	7	5	4	3	3		1		1	5		1	4		1	1	59
asp	6	6	2	4	4	7	7	2	1	3	5	5	2	3	1		2				60
glu	7	6	4	4	3	6	2	4	7	2		3		1	1	2	1	2			55
thr	6	6	6	3	5	3	3	2	3	3	7	3	2		4				1		57
lys	3	3	9	2	1	3	2	2	4	2	4	2	1	3	2		2	1			46
arg	4	1	4	4	5	1	2	1		3	2	2	2	3	3	2	2	1		1	43
pro	2	3	6	1	4	3	1	2	3		3	1	1	4	2	2	2				40
phe	2		5	2	3	2	5	2	2	2	3	2	2	1	1	1	1	2			38
cys	3	6	1	5	1		2	1	3	3			1	2	2	1		1	1		33
tyr	4		1	1	3	2	1	2	2	3	2		2		1	2	1	2			29
asn	3	3	2	1	2			3	1	4	2	1		3		5	1	1		1	33
gln	2	5	3	4	3	1			3	1	1	3	3			1			1	1	32
ileu	1	1	7		1	1	4	3	2	2	1	1	1		2		1	1			29
his	1			2	5		1	1	1		1	3	2				1		1		19
met	1						2		2						1				1		7
try		1	1	2	1			1	1												7
	80	77	73	69	69	60	52	47	56	43	41	38	37	37	31	30	25	20	8	7	900

Frequencies of dipeptides in 13 proteins and peptides. For this table $\chi^2 = 396{\cdot}2$, $n = 361$, $\sqrt{2\chi^2} - \sqrt{2n-1} = 1{\cdot}29$, which is less than the 5% confidence limit (1·65).

After successive additions of 100 dipeptides to the Table the number of different peptides present was counted (line A, Fig. 2) and for comparison 50 random sequences, each containing 4000 linkages, were drawn from a pool with the same composition as Table 5. The lines marked B in Fig. 2 show the limits within which 95% of the curves from the random sequences lie. The comparison suggests that chance is adequate to explain the absence of some dipeptides from Table 5 and that if dipeptides continue to be drawn randomly from the same pool a few types will probably not have been found even after the identification of four thousand linkages. The lack of bias

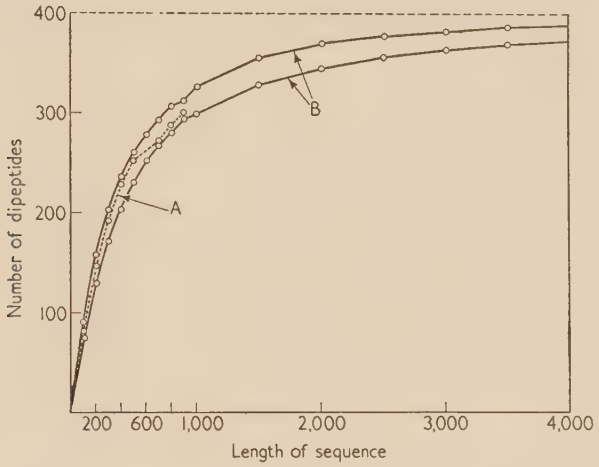


FIG. 2. Line A: Number of dipeptides represented in Table 5 after successive additions of 100 dipeptides.
Line B: Limits within which 95% of curves from random sequences fall.

in Table 5 is further supported by a value of 396·2 for χ^2 ($P > 0{\cdot}05$). Clearly the proteins and peptides used here do not support the conclusion of Needleman, Blackwell & Fosdick (1960) that "a natural barrier appears to exist for the occurrence of certain amino acid sequences in peptides."

5. Conclusion

In conclusion, it seems likely that opinions as to the reality of apparent patterns in amino acid sequences tend to ignore or underestimate the possible rôle of chance for, according to Cohen (1960), "nothing is so alien to the human mind as the idea of randomness." Chance cannot be estimated subjectively. Our purpose here has been to measure the frequency of chance patterns in sequences by an empirical method.

We thank Mr. H. P. F. Swinnerton-Dyer for his advice; Dr. F. Sanger, F.R.S. for his interest and the Director of the Mathematical Laboratory for permission to use the EDSAC 2. We are grateful to the Medical Research Council for financial support.

REFERENCES

Anderer, F. A., Uhlig, H., Weber, E. & Schramm, G. (1960). *Nature*, **186**, 922.
Anfinsen, C. B. (1959). In *The Molecular Basis of Evolution*. New York: John Wiley & Sons Inc.

Augenstine, L., Branson, H. R. & Carver, E. B. (1953). In *Information Theory in Biology*, ed. by H. Quastler, p. 105. Urbana: University of Illinois Press.
Bell, P. H. (1954). *J. Amer. Chem. Soc.* **76**, 5565.
Braunitzer, G., Hilschmann, N., Hilse, K., Liebold, B. & Müller, R. (1960). *Hoppe-Seyl. Z.* **322**, 96.
Braunitzer, G., Liebold, B., Müller, R. & Rudloff, V. (1960). *Hoppe-Seyl. Z.* **320**, 170.
Braunitzer, G., Rudloff, V., Hilse, K., Liebold, B. & Müller, R. (1960). *Hoppe-Seyl. Z.* **320**, 283.
Bromer, W. W., Sinn, L. G., Staub, A. & Behrens, O. K. (1956). *J. Amer. Chem. Soc.* **78**, 3859.
Cohen, J. (1960). In *Chance, Skill and Luck*. London: Penguin Books.
Folk, J. E. & Gladner, J. A. (1960). *Biochim. biophys. Acta*, **44**, 383.
Folk, J. E., Gladner, J. A. & Levin, Y. (1959). *J. Biol. Chem.* **234**, 2317.
Hirs, C. H. W., Moore, S. & Stein, W. H. (1960). *J. Biol. Chem.* **235**, 633.
Jollés, J. & Jollés, P. (1960). *Biochim. biophys. Acta*, **45**, 407.
Karadzova, M., Keil, B. & Šorm, F. (1960). *Coll. Czech. Chem. Commun.* **25**, 2878.
Keil, B., Šorm, F., Holeyšovský, V., Kostka, V., Meloun, B., Mikeš, O., Tomášek, V. & Vaněček, J. (1959). *Coll. Czech. Chem. Commun.* **24**, 3491.
Light, A. & Smith, E. (1960). *J. Biol. Chem.* **235**, 3144, 3151, 3159.
Lucas, F., Shaw, J. T. B. & Smith, S. G. (1958). *Advanc. Protein Chem.* **13**, 108.
Monier, R. & Jutisz, M. (1954). *Biochim. biophys. Acta*, **15**, 62.
Needleman, S. B., Blackwell, R. Q. & Fosdick, L. S. (1960). *Canad. J. Chem.* **38**, 477.
Pasynskiĭ, A. G. (1959). In *The Origin of Life on the Earth*, ed. by A. I. Oparin, A. G. Pasynskiĭ, A. E. Braunshtein & A. E. Pavlovskaya, p. 290. London: Pergamon Press.
Ryle, A. P., Sanger, F., Smith, L. F. & Kitai, R. (1955). *Biochem. J.* **60**, 541.
Skeggs, L. T., Lentz, K. E., Kahn, J. R., Shumway, N. P. & Woods, K. R. (1956). *J. Exp. Med.* **104**, 193.
Šorm, F. (1954). *Coll. Czech. Chem. Commun.* **19**, 1003.
Šorm, F. (1958). In *Symposium on Protein Structure*, ed. by A. Neuberger, p. 77. London: Methuen.
Šorm, F. (1959). In *The Origin of Life on the Earth*, ed. by A. I. Oparin, A. G. Pasynskiĭ, A. E. Braunshtein & A. E. Pavlovskaya, p. 231. London: Pergamon Press.
Šorm, F. & Keil, B. (1958). *Coll. Czech. Chem. Commun.* **23**, 1575.
Šorm, F., Keil, B., Holeyšovský, V., Knesslová, V., Kostka, V., Mäsiar, P., Meloun, B., Mikeš, O., Tomášek, V. & Vaněček, J. (1957). *Coll. Czech. Chem. Commun.* **22**, 1310.
Tuppy, H. (1958). In *Symposium on Protein Structure*, ed. by A. Neuberger, p. 66. London: Methuen.
Wilcox, P. E., Cohen, E. & Tan, W. (1957). *J. Biol. Chem.* **228**, 999.
Yčas, M. (1958). In *Information Theory in Biology*, ed. by H. P. Yockey, p. 70. London: Pergamon Press.

J. Mol. Biol. (1961) **3**, 541–546

The Molecular Structure of the Protein Matrix of Bovine Dental Enamel

Melvin J. Glimcher†, Laurence C. Bonar & Eubert J. Daniel

Department of Orthopedic Surgery, Harvard Medical School and the Massachusetts General Hospital, Boston, Massachusetts, U.S.A.‡

(*Received 17 March 1961*)

The protein matrix of bovine dental enamel, as well as various layers from the enamel-forming organ, remaining after decalcification with ethylenediaminetetra-acetic acid, have been isolated from mid-embryonic and early post-fetal incisor teeth. It has been possible to separate enamel matrix proper from the enamel organ. The enamel matrix proper gave a cross-β X-ray diffraction pattern. Furthermore, X-ray diffraction patterns of the enamel organ also indicated the presence of a cross-β component. These findings are discussed in relation to current concepts of the role of ordered organic matrices in biological mineralization.

1. Introduction

There has been increasing evidence that the molecular configuration and aggregation state of the structural proteins of mineralized tissues play a vital role in the initiation of calcification by a process of heterogeneous nucleation (Glimcher, Hodge & Schmitt, 1957; Watabe, Sharp & Wilbur, 1958; Neuman & Neuman, 1958; Strates & Neuman, 1958; Bachra & Sobel, 1959; Glimcher, 1959; Travis, 1960).

Most of the studies have utilized collagen as a model system in relation to the calcification of bone, cartilage, and dentin. However, dental enamel, the most highly mineralized of all mammalian tissues, with a mineral content of approximately 99% (Stack, 1955), has as its major organic constituent a protein whose amino acid composition is distinctly different from that of collagen (Losee & Hess, 1949; Block, Horwitt & Bolling, 1949; Eastoe, 1960). Electron microscopy has revealed that the organic matrix of enamel consists of thin extracellular filaments or fibrils which become rapidly impregnated with small inorganic mineral (hydroxyapatite) crystals (Watson, 1960; Frank, Sognnaes & Kern, 1960). This suggests that the organic matrix plays a role similar to collagen in inducing the formation of hydroxyapatite crystals. This paper will present data on the basic molecular configuration of the structural protein of microscopically identified cell-free bovine dental enamel.

2. Experimental

Embryonic calf heads (Holstein strain, male and female), ranging in age from 5 to 6 months, and newborn calf heads (Holstein and Angus strains, male and female) 2 to 10

† Fellow of the Medical Foundation of Boston.

‡ Communications should be addressed to the Orthopedic Research Laboratories, Massachusetts General Hospital, Fruit Street, Boston 14, Massachusetts.

days old, were procured from a slaughter house. The ages and sex of both embryos and newborn calves were determined by the abattoir veterinarian.†

Unerupted and erupted incisor teeth of the lower jaw were carefully dissected out, cleaned of surrounding connective tissue and capsular material, and rinsed briefly in cold buffer solution (0·02 M-veronal, KCl, $I = 0{\cdot}165$, pH 7·4). A silk suture was placed through the root of the larger teeth and the tooth was suspended crown downwards in individual vessels containing ethylenediaminetetra-acetic acid (EDTA). Smaller teeth were placed in capped bottles containing EDTA. Demineralization was carried out without stirring, but with careful changing of the EDTA, and concentrations varying from 0·05 M to 0·5 M and in pH from 7·0 to 8·5 were used. Demineralization was carried out at both 4°C and at room temperature with and without added preservative (thymol).

Within 2 to 10 days, depending on the age of the animal and the pH, EDTA concentration, volume and temperature of the decalcifying solution, soft tissue membranes were seen grossly. These were then separated from the tooth under the dissecting microscope. In embryonic teeth, two separate layers could be removed, the inner one of which was subsequently identified histologically as cell-free enamel matrix proper. Only a single layer, of varying histological composition, was evident on the teeth of newborn calves.

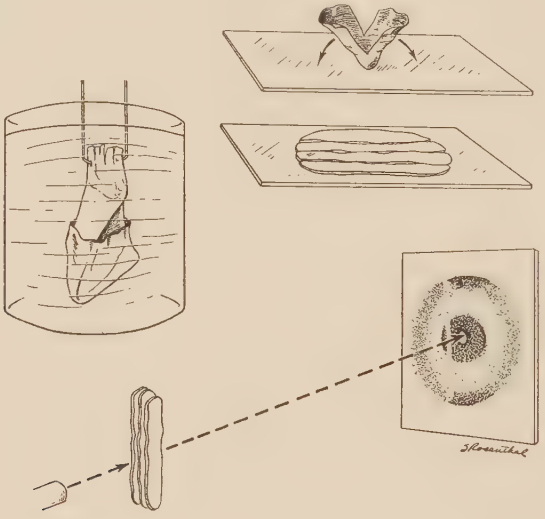


FIG. 1. Diagrammatic representation of the preparation and orientation of enamel matrix for X-ray diffraction. Preparation of the outer membrane is illustrated; the inner cell-free membrane could usually be removed from the buccal surface only, being too fragile for removal from the oral surface.

The membranes were fixed in 10% neutral formalin for 48 hr, dehydrated in alcohol and embedded in "Tissuemat" (m.p. 52·4°C). Sections 6 to 8 microns thick were cut on an International Minot rotary microtome, and stained by standard histological techniques.

Membranes were prepared for X-ray diffraction studies by washing free of buffer with distilled demineralized water, floating onto siliconized glass microscope slides and allowing to dry at room temperature. The dry membranes, usually about 10 to 20 μ thick, were removed from the slides with a razor blade. Small ($0{\cdot}3 \times 2{\cdot}0$ mm or less) pieces of the dried membrane were mounted directly, or, if several pieces were available, placed together and mounted; larger pieces of membrane were cut into strips approximately $0{\cdot}5 \times 2$ mm, several of which were placed together, and mounted.

X-ray diffraction patterns were obtained with a Phillips microcamera, specimen-to-film

† We gratefully acknowledge the assistance of Mr. Jack Snider of Snider Bros. Inc., meat packing house, and of Charles Gilvarg, D.V.M., U.S. Department of Agriculture, for their co-operation in securing and in determining the ages of the calves used.

distance 15 mm, equipped with a 100 μ bore glass capillary collimator. Copper Kα radiation, $\lambda = 1{\cdot}54$ Å, was obtained from a copper-target tube equipped with a nickel filter and operated at 35 kv. Specimens were mounted with the X-ray beam parallel to the membrane surface. Figure 1 illustrates the preparation and mounting of the specimens and the orientation of the specimens with respect to the X-ray beam.

For comparison, X-ray diffraction patterns were obtained from other bovine tissues of epidermal origin. The cornified layer of the gingiva (obtained by dissecting the transparent outer layer away from the underlying connective tissue of embryonic calf gingiva, previously softened by soaking for 72 hr in buffered isotonic KCl solution) was dried on a siliconized slide, and mounted as described above. Specimens of bovine hoof, horn and hair were mounted without further treatment. Lamellae of decalcified dentin, removed after stripping the tooth of enamel, were prepared for X-ray diffraction in a manner similar to the enamel matrix.

In addition, whole teeth were processed for histological examination as follows: undecalcified whole teeth were fixed in 10% neutral formalin for 48 hr, dehydrated in alcohol for 8 hr, and embedded in methyl methacrylate or in Araldite over a two-week period. Whole coronal sections, approximately 40 to 120 μ thick, were cut on a modified Gillings-Hamco sectioning apparatus, using a bonded diamond wheel 300 μ thick. These sections were examined in normal and polarized light.

Decalcified whole teeth, both prior to and after removal of the outer cellular layer and the enamel matrix proper were fixed, embedded in "Tissuemat" and coronal sections cut with the International Minot rotary microtome.

3. Results

Microscopic examination of whole coronal sections of incisor teeth from 5 to 6 month old embryos revealed the characteristic stellate reticulum, stratum intermedium, ameloblast layer and primary cuticle of the enamel organ as well as the enamel matrix proper, dentin and pulp (Plate I(a), (b)). Microscopic examination of the individual layers removed from embryonic teeth showed that the outer membrane consisted of elements of the enamel organ (Plate II(a)). The inner layer contained the cell-free enamel matrix proper (Plate II(b)). Examination of coronal sections of teeth, from which one and both membranes were removed, confirmed this observation. The single membrane removed from 2 to 10 day old calf incisors consisted of enamel matrix proper with varying amounts of residual enamel organ depending on the age of the animal and the area of the crown from which the membrane was obtained.

X-ray diffraction patterns of the enamel matrix proper showed an equatorial arc of spacing about 9·8 Å and a meridional arc at 4·65 Å, indicative of the cross-β structure (Astbury, Dickinson & Bailey, 1935). There was also a diffuse halo, of variable intensity relative to the 4·65 Å reflection, centered at about 4·6 to 5·1 Å. Diffraction patterns obtained from other tissues of epidermal origin (cornified outer layer of gingiva, hoof, hair, horn) showed characteristic α-keratin diffraction patterns, with no indication of the cross-β pattern. The dentin layer gave the distinctive collagen X-ray diffraction pattern. Typical diffraction patterns of the various tissues discussed above are shown in Plate III and the principal X-ray diffraction spacings for pertinent fibrous proteins are listed in Table 1 for comparison.

X-ray diffraction patterns obtained from the outer membrane of embryonic incisors, consisting of elements of the enamel organ, showed essentially the same features as the enamel matrix itself—the cross-β diffraction pattern and a diffuse halo.

Variations in the concentration or pH of the EDTA solutions, or the addition of thymol during demineralization, had no detectable effect on the X-ray diffraction patterns.

TABLE 1

Principal X-ray diffraction spacings of fibrous proteins (Å)

Protein	Meridional	Equatorial	Other
Collagen	2·86 s.	11–14 v.s. 5–6 s.	9·55 m; near-meridional 3·97 m; near-meridional 4·6 s; near-equatorial
α-keratin	5·1 s.	9·8 s.	
β-keratin	3·33 w.–m.	9·7 s. 4·65 s.	
Cross-β-keratin	4·6–4·7 s.	9·8 s.	
Protein matrix of bovine dental enamel	4·65 ± 0·05 s.	9·8–10·0 s.	4·6–5·1 m.–s. diffuse halo

v.s. = very strong; s. = strong; m.s. = moderately strong; m. = moderate; w. = weak.

4. Discussion

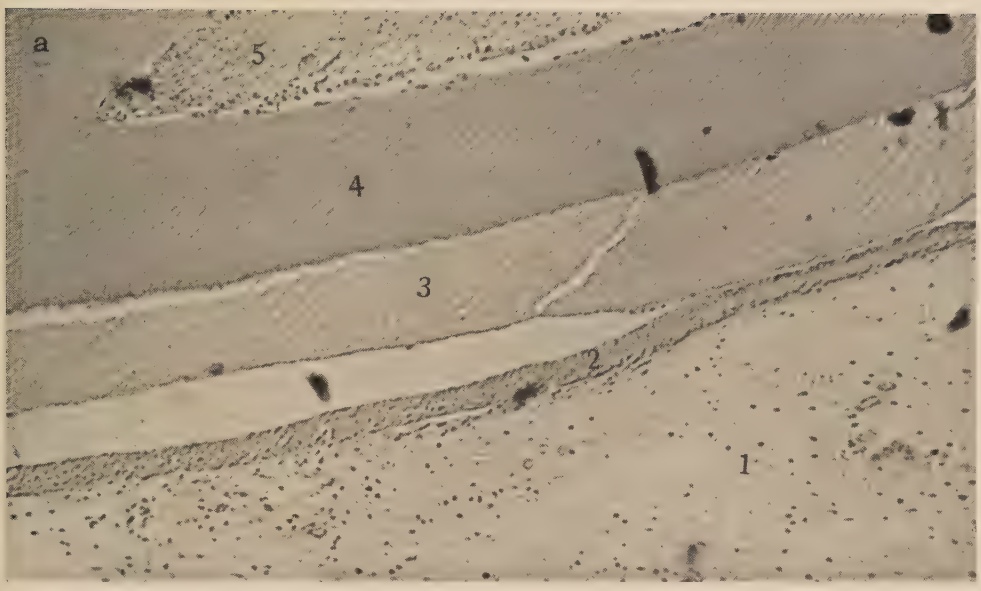
The structural protein of cell-free bovine dental enamel matrix proper which remains after demineralization with EDTA is sufficiently ordered to give a fiber diffraction pattern. Its cross-β molecular configuration distinguishes it from the more common α- and β-keratins and from collagen. This configuration appears to be peculiar to the dental enamel, since other bovine keratins of epidermal origin (hoof, hair and horn), as well as the gingival epithelial layer adjacent to the teeth, all gave the α-keratin diffraction pattern.

The cross-β X-ray diffraction pattern obtained from the membrane-containing elements of the enamel organ is presumably due to the cross-β component being present either intracellularly or in the acellular primary cuticle.

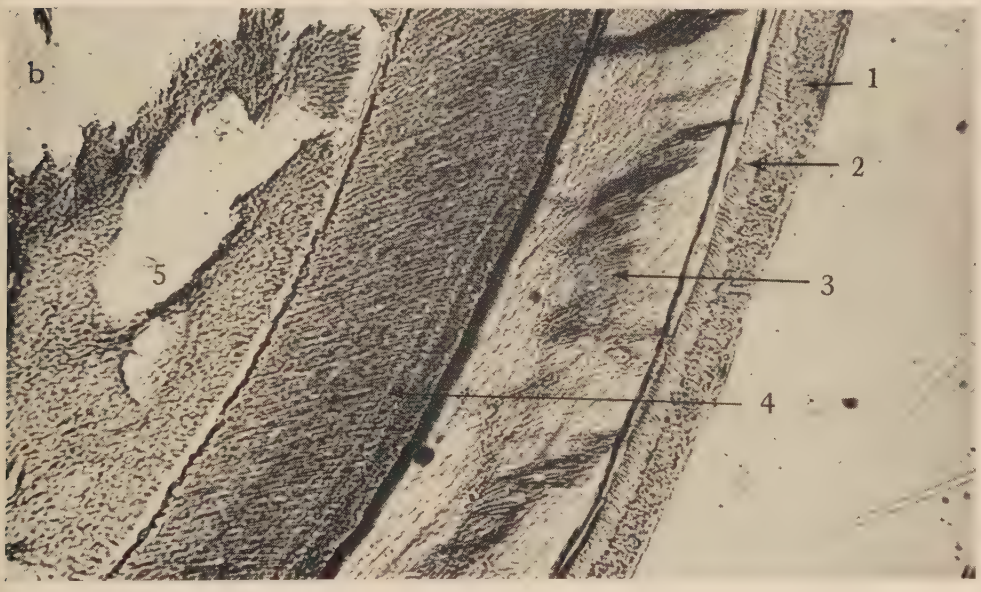
The existence of a diffuse halo at 4·6 to 5·1 Å indicates the presence of a considerable amount of relatively unordered protein. This may be a second, unordered component, such as exists in feather keratin (Fraser & McRae, 1959) or a disordered portion of the protein producing the cross-β diffraction pattern.

The only previously reported naturally-occurring cross-β proteins are the silk-like protein of the *Chrysopa* egg stalk (Parker & Rudall, 1957) and the protein of bacterial flagella (Astbury, Beighton & Weibull, 1955). The cross-β structure has usually been noted in proteins subjected to decidedly non-physiological conditions, involving heating (Astbury *et al.*, 1935; Rudall, 1952) or chemical scission of cross links (Peacock, 1959). These conditions cause the transformation of the α-protein into the cross-β modification. Pautard (1958, 1959) has reported that the treatment of actomyosin films with ATP results in the appearance of a cross-β component. Since the preparative methods employed here are of a relatively mild nature, and would not be expected to induce a transformation of molecular configuration, it is concluded that a portion of the protein matrix of enamel is in the cross-β configuration *in vivo*.

The nature of the polypeptide chain configuration responsible for the cross-β diffraction pattern is still under investigation (Astbury, Beighton & Parker, 1959). Although it has been suggested (Zubay, 1959) that the cross-β diffraction pattern is



(a)



(b)

PLATE 1(a). Coronal section of a paraffin-embedded decalcified five-month embryonic calf incisor tooth. 1, Stellate reticulum and stratum intermedium; 2, ameloblast layer; 3, enamel matrix proper; 4, dentin; 5, pulp cavity. ×100.

Note the separation of an outer layer consisting of stellate reticulum and ameloblasts, and an inner layer consisting of the enamel matrix proper.

(b) Coronal section of a plastic embedded undecalcified five-month embryonic calf incisor tooth seen with partially polarized light, showing the same layers as in (a). ×145.

[*To face page* 544

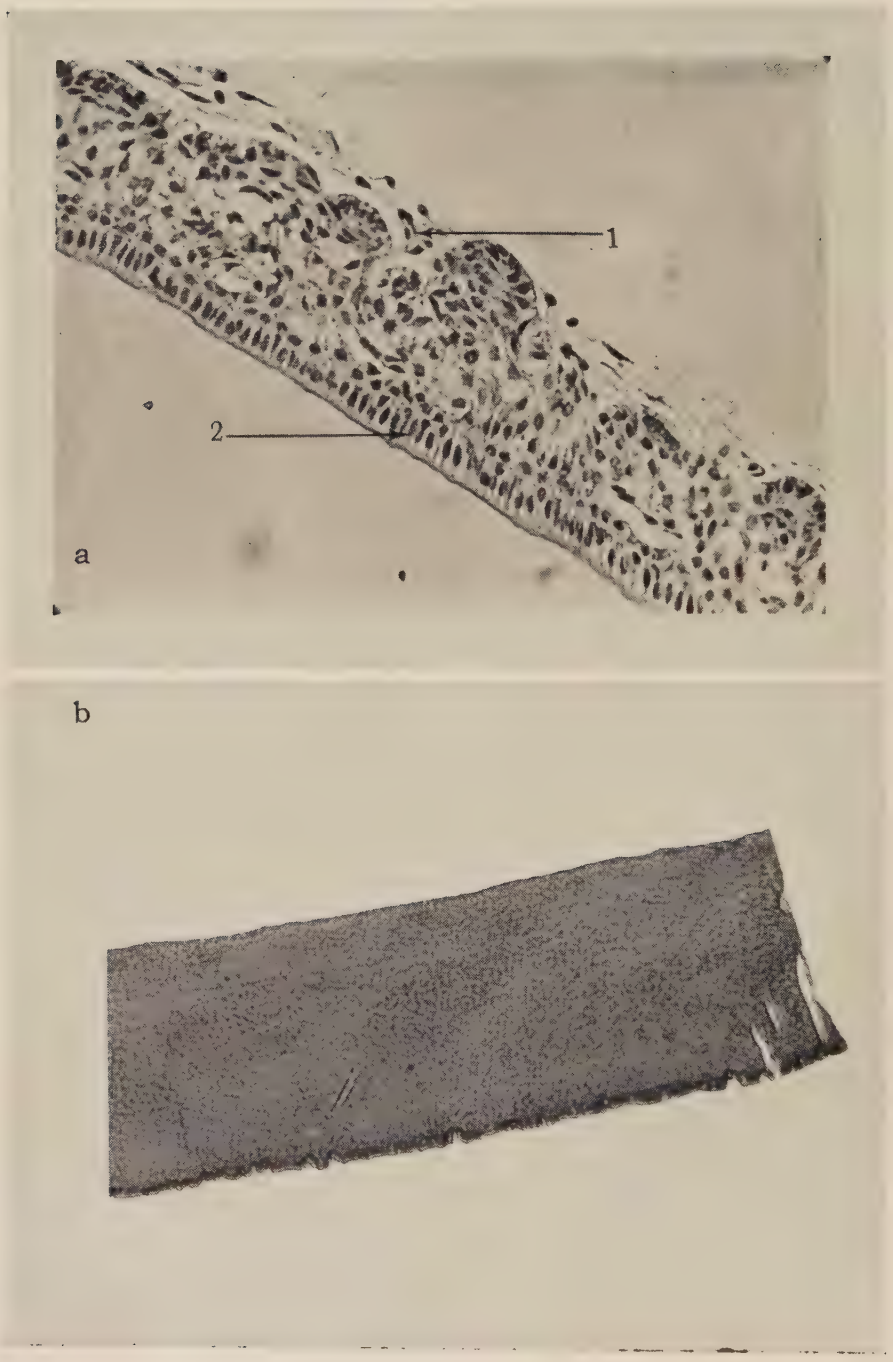
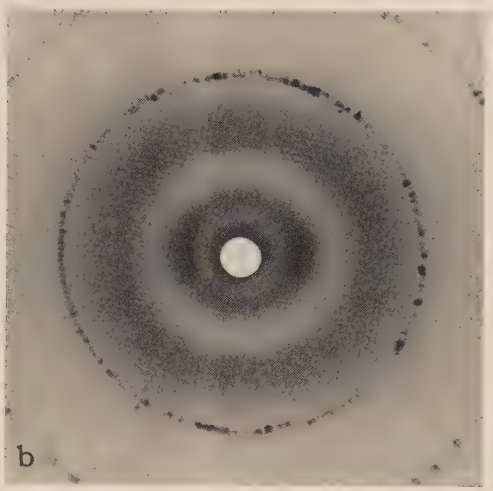


PLATE II(a). Cross section of the outer membrane removed from decalcified five-month embryonic calf incisor tooth. Layers 1 and 2 as in Plate I(a). ×330.

(b) Cross-section of the inner membrane removed from decalcified five-month embryonic calf incisor tooth consisting of the cell-free enamel matrix proper. ×143.



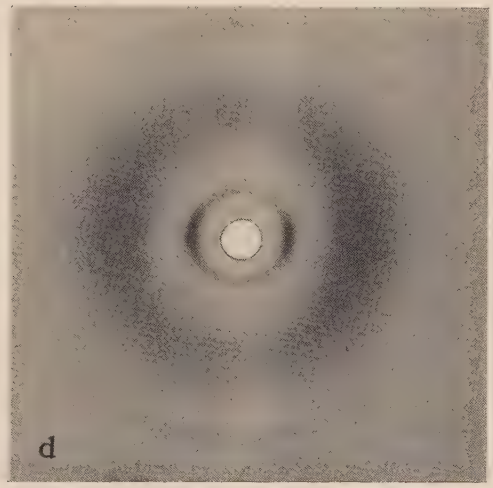


PLATE III. X-ray diffraction patterns.

(a) Enamel matrix proper of new-born calf incisor tooth showing the 4·65 Å meridional reflection characteristic of the cross-β-pattern.

(b) Cornified layer of gingiva, a poorly defined α-keratin pattern typical of epidermal keratin. Salt ring due to KCl.

(c) Calf hair, a well oriented α-keratin pattern.

(d) Dentin, exhibiting characteristic collagen pattern.

(The streak running through the center of some patterns is the shadow of the specimen, which is considerably thicker parallel to the X-ray beam than laterally perpendicular to it.)

due to parallel α-helices tilted approximately 20° from the fiber axis, it seems probable, on the basis of X-ray diffraction and infrared spectrum studies (Astbury *et al.*, 1959; Parker & Rudall, 1957; Ambrose & Elliott, 1951), that the polypeptide chains in cross-β proteins run predominantly transverse to the fiber axis, with the C=O and N—H bonds parallel to the fiber axis, and the amino acid side-chains approximately perpendicular to the fiber axis. Recently Bradbury *et al.* (1960) have prepared films of poly-β-*n*-propyl-L-aspartate, poly-β-benzyl-L-aspartate, and poly-γ-benzyl-L-glutamate of low degree of polymerization in the cross-β configuration, as shown by infrared spectra and X-ray diffraction patterns.

If the polypeptide chain configuration and pattern of hydrogen bonding in the cross-β and parallel-β structures are similar, resembling one of the pleated-sheet structures proposed by Pauling & Corey (1951), two alternative cross-β structures are conceivable. The axis of the polypeptide chains may be linear and transverse to the fiber axis, in which case the cross-β structure differs from the parallel-β only in that the fiber (i.e. rotation) axis is parallel to a different unit cell axis (Bradbury *et al.*, 1960); or the polypeptide chains may run transverse to the fiber axis for relatively short distances, then be folded back on themselves as suggested by Astbury *et al.* (1959).

Recent electron microscopic studies of dental enamel (Frank *et al.*, 1960) have indicated that the organic matrix of rat and human enamel is in the form of filaments of mean diameter about 50 Å, approximately perpendicular to the tooth surface, with lateral bridges connecting these filaments. Because of the marked (40-fold or more) decrease in thickness of the enamel layers on drying, caution must be exercised in interpreting the orientation indicated by the X-ray diffraction patterns. If the length of the well-ordered regions of the filaments is small compared to the thickness of the membrane and shrinking occurs mainly in the intervening unordered regions, then the membranes can shrink without change in the orientation of the polypeptide chains with respect to the membrane surface. On the other hand, if the filaments collapse (either accordion-fashion or like up-ended matches) the filament axes will lie parallel to the surface of the dried membrane. In the former case, the cross-β fiber axis is perpendicular to the filament axis in intact enamel, and the polypeptide chains are approximately parallel to the axis. In the latter case the cross-β fiber axis is parallel to the filament axis, and the polypeptide chains run predominantly perpendicular to the filament axis. The exact geometrical relationship between filament axis and polypeptide chain direction must obviously await further X-ray diffraction studies on undried enamel matrix, correlated with electron microscopic studies.

Regardless of the relative orientation of fiber axis and filament axis, the orientation of the 4·65 Å reflection perpendicular to the 9·6 Å reflection indicates that the polypeptide chains are in the cross-β rather than the parallel-β configuration; the cross-β structure differs fundamentally from the parallel-β in that the rotation axis is parallel to the C=O . . . . . H—N hydrogen bonds, rather than to the polypeptide chain axis, as in the parallel-β case.

The present X-ray diffraction results confirm the conclusions of previous chemical analyses that the enamel protein is distinct from collagen. This suggests that a protein other than collagen, with a different but nevertheless well-ordered molecular structure, may be capable of initiating the deposition of apatite crystals. If this were the case, it would lend further support to the concept of the role of molecularly ordered organic matrices in biological mineralization (Glimcher, 1960).

The authors are indebted to Dr. Philip T. Levine of Tufts University School of Dental Medicine for valuable suggestions and assistance in preparing the dental enamel, and to Dr. Paul Goldhaber of the Harvard School of Dental Medicine for helpful discussion during the course of this work.

This work has been supported by grants from the United States Public Health Service (Research Grant No. 6391-Multiple), the John A. Hartford Foundation, Inc., and The Easter Seal Research Foundation of the National Society for Crippled Children and Adults, Inc.

REFERENCES

Ambrose, E. J. & Elliott, A. (1951). *Proc. Roy. Soc.* A, **208**, 75.
Astbury, W. T., Beighton, E. & Parker, K. D. (1959). *Biochim. biophys. Acta*, **35**, 17.
Astbury, W. T., Beighton, E. & Weibull, C. (1955). *Proc. Symp. Soc. Exp. Biol.* **9**, 282.
Astbury, W. T., Dickinson, S. & Bailey, K. (1935). *Biochem. J.* **29**, 2351.
Bachra, B. N. & Sobel, A. E. (1959). *Arch. Biochem. Biophys.* **85**, 9.
Block, R. J., Horwitt, M. K. & Bolling, D. (1949). *J. Dent. Res.* **28**, 518.
Bradbury, E. M., Brown, L., Downie, A. R., Elliot, A., Fraser, R. D. B., Hanby, W. E. & McDonald, T. R. R. (1960). *J. Mol. Biol.* **2**, 276.
Eastoe, J. E. (1960). *Nature*, **187**, 411.
Frank, R. M., Sognnaes, R. F. & Kearn, R. (1960). *Amer. Ass. Advanc. Sci., Wash.* No. 64, 163.
Fraser, R. D. B. & MacRae, T. P. (1959). *J. Mol. Biol.* **1**, 387.
Glimcher, M. J. (1959). *Rev. Mod. Phys.* **31**, 359.
Glimcher, M. J. (1960). *Amer. Ass. Advanc. Sci., Wash.* No. 64, 421.
Glimcher, M. J., Hodge, A. J. & Schmitt, F. O. (1957). *Proc. Nat. Acad. Sci., Wash.* **43**, 860.
Losee, F. L. & Hess, W. C. (1949). *J. Dent. Res.* **28**, 512.
Neuman, W. F. & Neuman, M. W. (1958). In *The Chemical Dynamics of Bone Mineral.* Chicago Ill.: University of Chicago Press.
Parker, K. D. & Rudall, K. M. (1957). *Nature*, **179**, 905.
Pauling, L. & Corey, R. B. (1951). *Proc. Nat. Acad. Sci., Wash.* **37**, 729.
Pautard, F. G. E. (1958). *Nature*, **182**, 788.
Pautard, F. G. E. (1959). *Nature*, **183**, 1391.
Peacock, N. (1959). *Biochim. biophys. Acta*, **32**, 220.
Rudall, K. M. (1952). *Advanc. Prot. Chem.* **7**, 253.
Stack, M. V. (1955). *Ann. N.Y. Acad. Sci.* **60**, 585.
Strates, B. & Neuman, W. F. (1958). *Proc. Soc. Exp. Biol. Med.* **97**, 688.
Travis, D. F. (1960). *Amer. Ass. Advanc. Sci., Wash.* No. 64, 57.
Watabe, N., Sharp, D. G. & Wilbur, K. M. (1958). *J. Biophys. Biochem. Cytol.* **4**, 281.
Watson, M. L. (1960). *J. Biophys. Biochem. Cytol.* **7**, 489.
Zubay, G. (1959). *J. Mol. Biol.* **1**, 186.

J. Mol. Biol. (1961) **3**, 547–565

The Molecular Configuration of Deoxyribonucleic Acid III. X-ray Diffraction Study of the *C* form of the Lithium Salt

D. A. Marvin

Wheatstone Laboratory, King's College, London, England

M. Spencer, M. H. F. Wilkins

Medical Research Council Biophysics Research Unit, King's College, London, England

AND

L. D. Hamilton

Sloan-Kettering Institution, New York, U.S.A.

(*Received 17 March 1961*)

The lithium salt of deoxyribonucleic acid may have either of two distinct molecular configurations, *B* or *C*. The *B* and *C* X-ray diffraction patterns are very similar, but these are nevertheless distinct. The distinction is accounted for by slight differences between the configurations. Depending on conditions in the fibres, molecules in the *C* form may pack in either a hexagonal or an orthorhombic lattice; these conditions (e.g. water content and salt content) are discussed. Techniques used in analysing partly-disordered diffraction patterns are described. The *C* configuration consists of a non-integral helix with about $9\frac{1}{3}$ residues in a pitch of 31 Å; the base pairs are moved from their positions in the *B* model by 1·5 Å in such a direction that the narrow groove is made deeper, and are tilted by 6°. The configuration of a single nucleotide is essentially the same as in *B*: only the position and orientation of the nucleotides within the helix have been altered. Packing of smooth double helices in orthorhombic and hexagonal lattices is discussed, and the conclusions are applied to the packing of *C* molecules.

1. Introduction

X-ray diffraction is the only method available for precise determination of the three-dimensional configuration of DNA. Diffraction data from fibres of DNA are less extensive than data from the single crystals normally used in molecular structure determination. Therefore in studying the DNA structure, it is desirable to confirm results by analysing several different types of diffraction pattern, each corresponding to a slightly different configuration of the molecule. Moreover, it would be possible to establish with greater certainty that a postulated structure was unique if one could distort the DNA molecule and show that the diffraction pattern changed in the expected manner. This is achieved by the comparison, described in this paper, of the *C* and *B* configurations of DNA: the *C* form may be regarded as a distorted *B* configuration.

Previous papers in this series have described in detail the determination of the *B* configuration of DNA. Part I (Langridge *et al.*, 1960*a*) gave experimental techniques and X-ray diffraction data from a crystalline *B* form of the lithium salt of DNA (LiDNA). Part II (Langridge *et al.*, 1960*b*) described the techniques of building and refining molecular models to agree with the diffraction data. In the present paper these techniques are applied to determining the *C* configuration (Marvin, Spencer, Wilkins & Hamilton, 1958) of LiDNA. Details which are the same as in Part I will not be described again; new techniques demanded by the different nature of the problem will be given.

2. Nature of the Diffraction Data

Diffraction patterns obtained from polymer fibres have several characteristics. If the molecules are parallel in the fibre and consist of identical units regularly repeated along their length, diffracted intensity is confined to layer lines perpendicular to the fibre direction. If the molecules are laterally packed in regular crystalline array, diffracted intensity appears as discrete sharp reflexions ("spots") on the layer lines. Since the crystallites within the fibre are generally aligned not exactly parallel, these spots are drawn out into arcs. (See James, 1954; Stokes, 1955.)

The crystalline *B* form of DNA gives a diffraction pattern consisting entirely of sharp spots (Part I). The *C* form, on the other hand, gives a pattern with spots in the centre and diffuse intensity distributions ("streaks") in the outer part of the pattern. The spot-size increases with increasing diffraction angle and in the region of 10 Å the spots merge into the streaks. This type of pattern, with spots in the centre and streaks elsewhere, we call "semicrystalline." The spots on the *C* patterns occur in either of two different arrangements (Plates I(a) and (b)) corresponding to either of two kinds of molecular packing. However, the overall intensity distribution is the same on the two patterns; this indicates that the configuration of the individual molecules is the same in the two cases.

The diffuseness in semicrystalline fibre patterns of helical molecules may be caused by various types of disorder in the fibres:

(1) Molecules may be randomly displaced from regular positions in the crystallite. Reflexions are not broadened, but fade with increasing diffraction angle, merging into diffuse intensity regions (Hosemann, 1951; Amoros, Canut & De Acha, 1960).
(2) The molecules may be randomly screwed along the helix direction. Such screw disorder (see section 6) gives rise to spots in the central ($m = 0$) region of the pattern and layer line streaks elsewhere (Klug, Crick & Wyckoff, 1958).
(3) Individual crystallites may be distorted. Reflexions are broadened roughly in proportion to diffraction angle (Stokes & Wilson, 1944).
(4) Different crystallites in the fibre may have different lattice parameters. Again, broadening of reflexions increases with angle.
(5) Crystallites may be so small that diffraction broadening of reflexions occurs. All reflexions are broadened by effectively the same amount.

In practice these effects occurred in varying proportions on different patterns. For instance Plate I(a) shows sharper central reflexions than Plate I(b): factors (1) and (2) may be contributing more in I(a) than in I(b).

In earlier work (Wilkins, Stokes & Wilson, 1953), disordered *B* patterns similar to the semicrystalline *C* have been referred to as "paracrystalline". These early patterns

were ill-defined, so it was not clear what sort of disorder was present. We now think the term "paracrystalline" (Hosemann, 1950*a*,*b*) does not fully cover the types of disorder in DNA fibres; we prefer the more general "semicrystalline" to distinguish disordered DNA patterns from fully crystalline patterns.

3. The Two Types of *C* Diffraction Pattern

(a) *Determination of the lattice parameters*

Reciprocal space distances were obtained from the diffraction patterns as described in Part I. Only rough lattice parameters could be determined from the few spots on the diffraction patterns. Preliminary indexing suggested that one of the two types of pattern fitted a hexagonal lattice, and the other an orthorhombic lattice (Plate II). Unit cell dimensions in the *ab* plane indicated that three molecules passed through the hexagonal unit cell, and two molecules through the orthorhombic; the height of the unit cell in the *c* direction, given by the layer line spacing, was one turn of the helix. Both the hexagonal group of three and the orthorhombic group of two molecules have been observed for the semicrystalline *B* form of DNA (Table 1 of Part I).

Consideration of systematically absent reflexions on the orthorhombic diffraction patterns showed that the relative positions of the molecules in fractions of the unit cell edge were (0,0,0,) and ($\frac{1}{2}$, $\frac{1}{2}$, $\frac{1}{4}$). There may be up to 10% error in the z value, but comparison of Plate II(a) with Fig. 1(a) shows that the relative displacement is closer to $\frac{1}{4}$, than $\frac{1}{3}$, e.g. the 112 and 202 reflexions are weak and 213 is not absent. On the hexagonal patterns, systematic absences indicated that relative positions of the molecules were (0,0,0), ($\frac{1}{3}$, $\frac{2}{3}$, z) and ($\frac{2}{3}$, $\frac{1}{3}$, $-z$), where z is either $\frac{1}{6}$ or $\frac{1}{3}$. The hexagonal patterns showed streaks on the 1st and 3rd layer lines even in the centre of the pattern, indicating a tendency for the molecules to be randomly translated by $\pm\frac{1}{2}$ in the z direction (Feughelman *et al.*, 1955; Wyckoff, 1955; Klug & Franklin, 1958). Comparison of faint spots superimposed on the 1st and 3rd layer line streaks of Plate II(b) with Fig. 1(b) indicated that in ordered regions of the fibre the relative z coordinates of the three molecules are probably 0, $\frac{1}{6}$, $-\frac{1}{6}$, and not 0, $\frac{1}{3}$, $-\frac{1}{3}$; however, with a completely random z translation of $\frac{1}{2}$, there would be no distinction between these two schemes.

For more detailed parameter determination, the lattice was assumed to be precisely hexagonal or orthorhombic, and spots were then fitted to the best possible lattice parameters a,b,c, either by a least squares procedure or by trial-and-error. The strong meridional reflexion (all fibres were tilted by 13°30′ to observe this reflexion) was measured in order to obtain the separation of residues along the helix. Values obtained from measurement of some of the best *C* patterns are shown in Table 1. The average values for the two hexagonal patterns used in intensity measurements are $a = 35{\cdot}0$ Å, $c = 30{\cdot}8$ Å, base separation $p = 3{\cdot}32$ Å; for the two orthorhombic patterns used, $a = 32{\cdot}2$ Å, $b = 20{\cdot}2$ Å, $c = 31{\cdot}1$ Å, $p = 3{\cdot}32$ Å. For the purpose of molecular model building, a value of 3·32 Å was taken for the nucleotide separation; the helix was assumed to have exactly $9\frac{1}{3}$ residues per turn; the pitch length was 31·0 Å.

(b) *Factors determining molecular packing*

Franklin & Gosling (1953*a*) have described the dependence of the sodium DNA diffraction pattern on relative humidity (r.h.). At about 75% r.h. the crystalline *A* form (Franklin & Gosling, 1953*b*, 1955; Langridge *et al.*, 1957) may be obtained; at

higher humidities the semicrystalline *B* form (Wilkins *et al.*, 1953; Feughelman *et al.*, 1955) is obtained. In the case of LiDNA, at 66% r.h. the orthorhombic crystalline *B* form may be obtained; at higher humidities the semicrystalline *B* form is obtained (Part I).

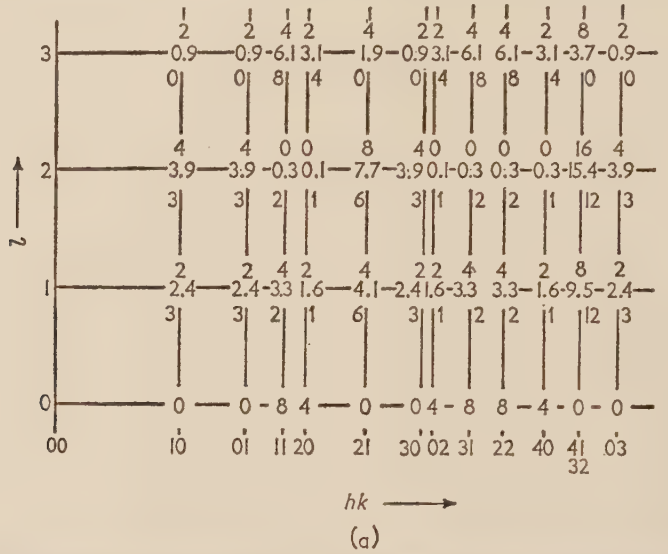
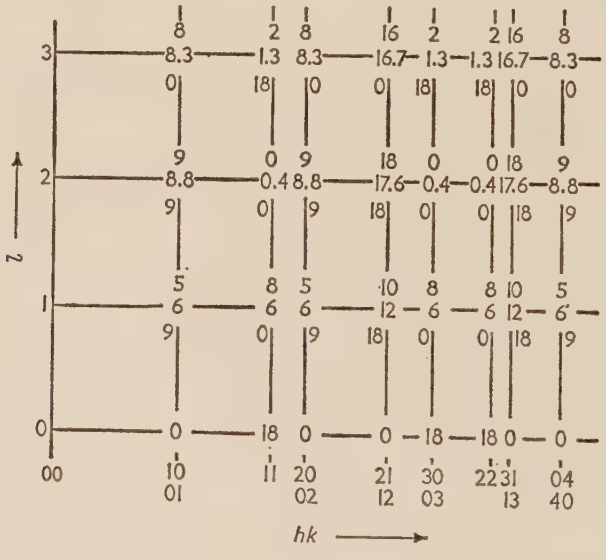


FIG. 1. Row lines and relative intensities in a fibre diagram of (a) orthorhombic group of two molecules at (0, 0, 0,) and ($\frac{1}{2}$, $\frac{1}{2}$, z). The top figure represents relative intensity for $z = \frac{1}{4}$, the middle for $z = 0{\cdot}28$, and the bottom for $z = \frac{1}{3}$.



(b) hexagonal group of three molecules at (0, 0, 0), ($\frac{1}{3}$, $\frac{2}{3}$, z), and ($\frac{2}{3}$, $\frac{1}{3}$, $-z$). The top figure represents relative intensity for $z = \frac{1}{6}$, the middle for $z = 0{\cdot}19$, and the bottom for $z = \frac{1}{3}$.

We have found for LiDNA that the type of diffraction pattern depends on the concentration of LiCl in the specimen, as well as on the water content (Marvin, 1960). Crystalline *B* diffraction patterns are obtained when 1 to 6% by weight of

PLATE I(a). Orthorhombic *C* diffraction pattern taken at 44% r.h.; fibre tilted by 13·5° from perpendicular to the incident beam. The same fibre at 66% r.h. gave Plate IV of Part I.

[*To face page* 550

PLATE I(b). Hexagonal *C* pattern; 66% r.h.; 13·5° tilt. Note spots on equator and 2nd layer lines and streaks on 1st and 3rd layer lines. Diffuse diffraction on the meridian is due to white X-rays.

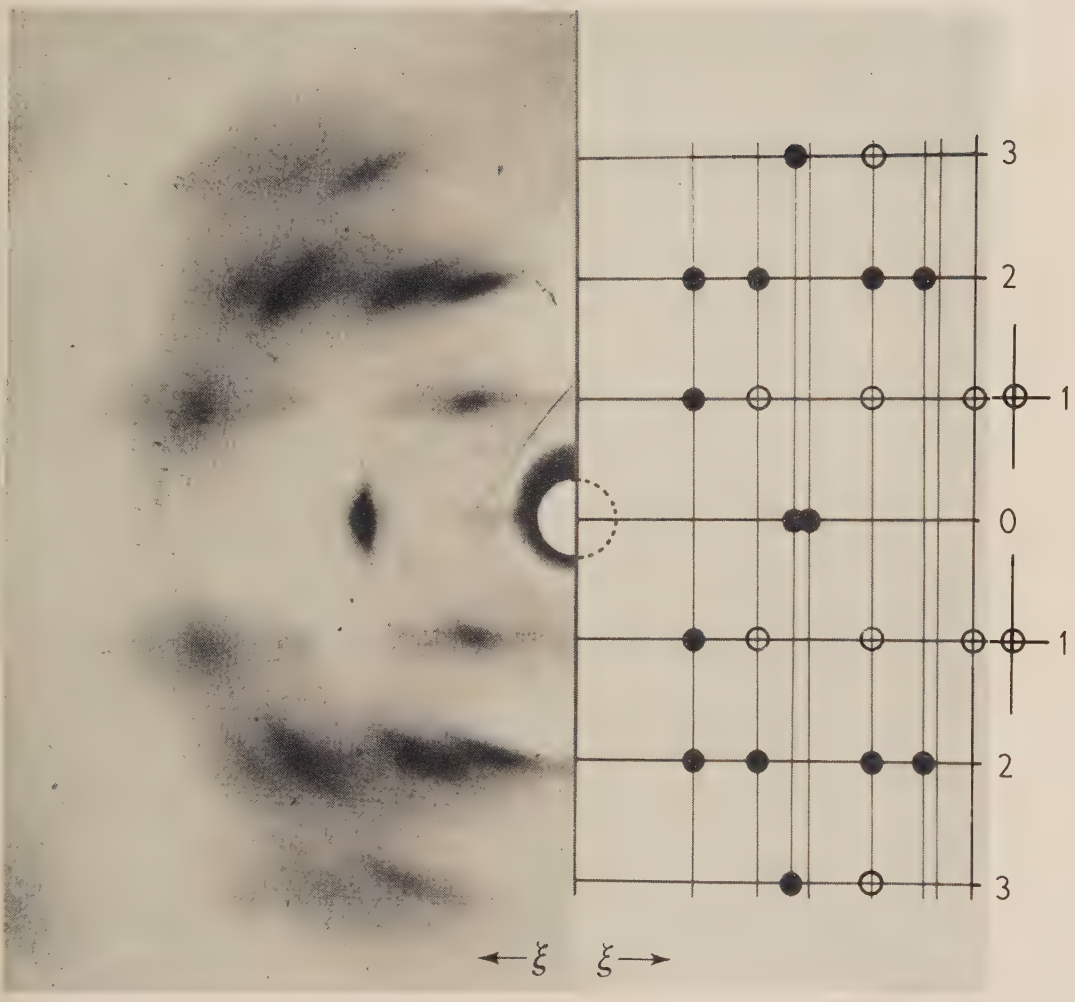


PLATE II(a). Central part of Plate I(a) showing arrangement of row lines for an orthorhombic lattice and observed spots in the centre of pattern. Measured spots shown ●, less well defined spots shown ○.

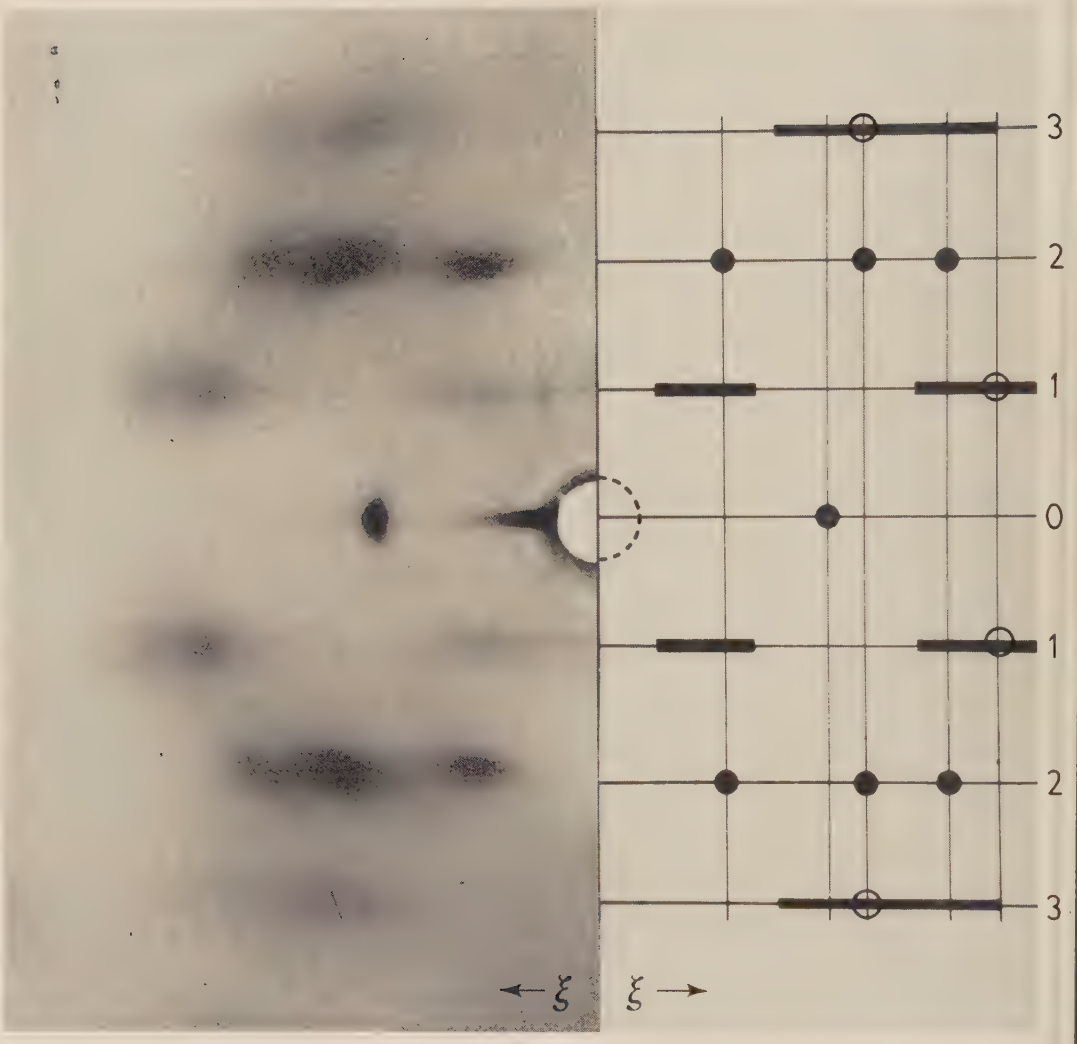


PLATE II(b). Central part of Plate I(b) showing row lines for a hexagonal lattice. There are spots on even layer lines and streaks on odd layer lines. On odd layer lines there are some faint indications of spots shown C
In both (a) and (b) the meridian of the diffraction pattern coincides with that of the lattice diagram.

PLATE III(a). Hexagonal *C* pattern, taken over P_2O_5, 13·5° tilt. The same fibre previously gave Plate I(a). ote absence of the equatorial spot.

PLATE III(b). Hexagonal *C* pattern at 44% r.h., 13·5° tilt. Same fibre as III(a). The lattice has expanded so that the equatorial spot missing in (a) is once more visible.

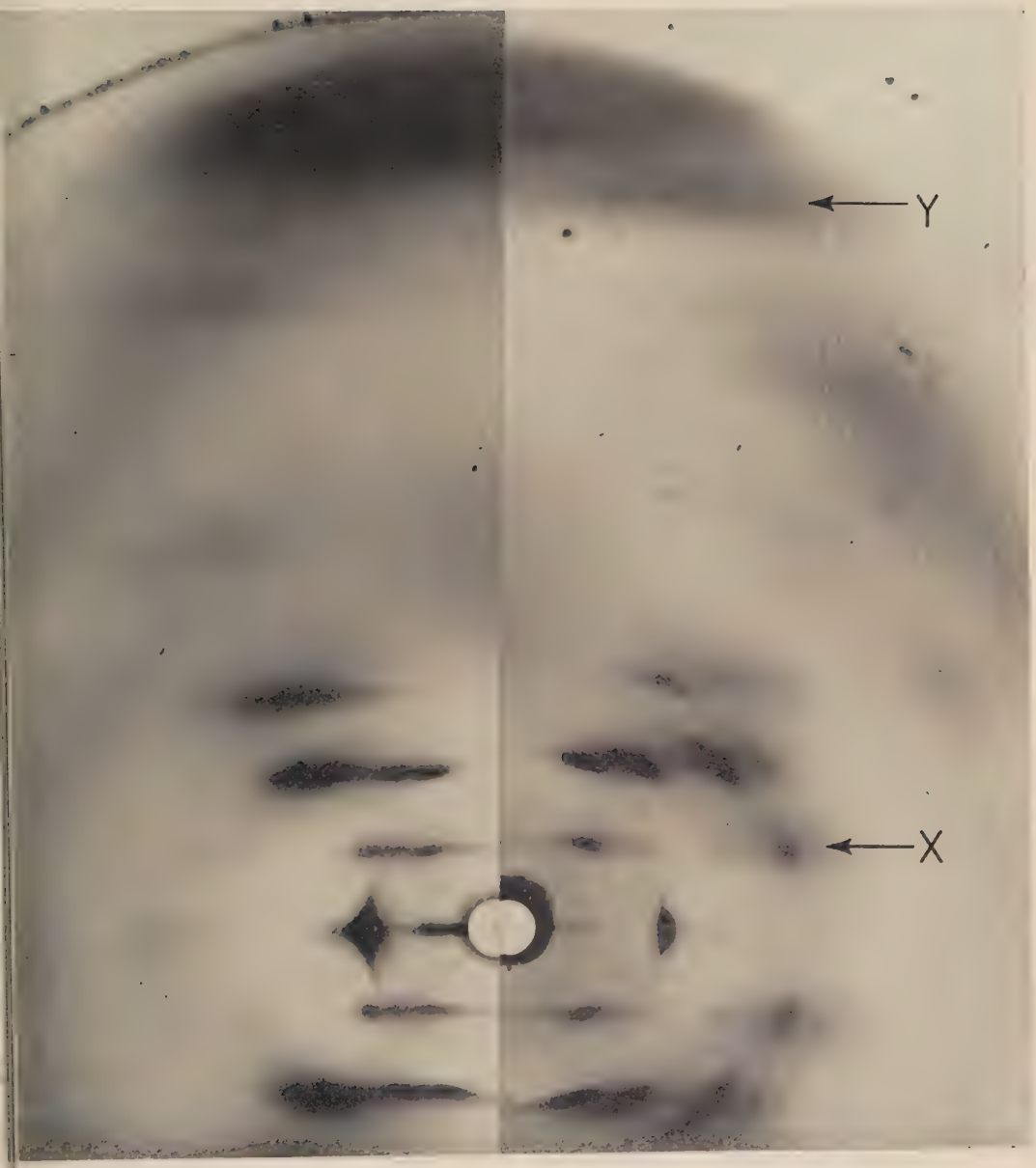


PLATE IV. Comparison of semicrystalline orthorhombic *B* (left) and *C* (right) diffraction patterns. At X note strong intensity for $\xi = 0{\cdot}10$ Å^{-1} on the 1st layer line of *C* and corresponding weak intensity on *B*. At Y note the strong J_1 term from the $m = 1$ system of Bessel functions on *C* and corresponding weak intensity on *B*.

chloride (as LiCl) has been precipitated with the DNA during preparation; the optimum value is 2·5 to 4%. Specimens with less than 1% chloride give only *C* or semicrystalline *B* patterns; specimens with much more than 6% chloride either give high humidity *B* patterns or are too deliquescent to be used. The chloride measurements were taken at 50 to 60% r.h., so about 25% by weight of the specimen was water (Part I); a micro-Carius method (e.g. Pregl, 1951) was used to determine chloride. Experiments were also performed in which 0·01 M to 0·2 M-LiCl solutions were added to small quantities of air-dried, chloride-free LiDNA, in amounts to give the above concentrations of chloride. Such specimens gave crystalline patterns, but better patterns were obtained from specimens in which the LiCl was precipitated with the DNA during its preparation. The optimum chloride content corresponds to about one LiCl molecule per two nucleotides; the range over which crystallinity could be obtained corresponds to between one LiCl per ten nucleotides and one LiCl per nucleotide. Similar experiments on the double-stranded synthetic polyribonucleotide polyinosinic + polycytidylic acid (Davies, 1960) indicate that crystallinity of this polymer is induced by the presence of about one sodium cacodylate molecule per two nucleotides.

If chloride-containing fibres which have given crystalline *B* patterns at 66% r.h. are brought immediately to a lower humidity, the orthorhombic *C* pattern is always obtained. If fibres with low chloride content are tested at 66% r.h. or below, they give the *C* pattern, usually in the hexagonal lattice. All fibres give semicrystalline *B* patterns at 92% r.h. The humidity at which one type of pattern changes to another varies with the fibre. In addition there are hysteresis effects which affect the type of pattern obtained. For instance a fibre containing 2·8% chloride gave a crystalline *B*

TABLE 1

Lattice parameters for orthorhombic and hexagonal C *patterns at various relative humidities*

Fibre	Type	r.h.%	*a*	*b*	*c*	*p*	*c*/*p*	
A	hexag.	57	32·1		30·4	3·34	9·1	
B	hexag.	57	35·4		30·9	3·32	9·3 *	
C	ortho.	44	32·4	20·6	31·2	3·32	9·4 *	
D	ortho.	32	31·2	20·0	30·8	3·32	9·3	
E	hexag.	66	34·5		30·7	3·32	9·3 *	(Plate I(b))
	gave semicrystalline orthorhombic *B* at 92% r.h.; then gave							
	ortho.	57	31·6	19·8	31·2	3·28	9·5	
F	ortho.	44	32·0	19·8	31·0	3·32	9·3 *	(Plate I(a))
	gave semicrystalline hexagonal *B* at 92% r.h.; then dried 6 hours over P_2O_5 and then gave successively							
	hexag.	dry	30·8		29·2	3·31	8·8	
	hexag.	32	33·0		29·4	3·31	8·9	
	hexag.	44	34·0		29·9	3·31	9·0	(Plate III(b))
	hexag.	75	34·7		32·2	3·31	9·7	
	hexag.	57	33·0		30·8	3·31	9.3	

Hysteresis effects are apparent in fibres E and F. Patterns marked * were used for intensity measurements. (The salts used for obtaining 32%, 44%, 57%, 66%, and 75% r.h. were $CaCl_2$, K_2CO_3, NaBr, $NaNO_2$ and $NaClO_3$, respectively. "Dry" patterns were obtained with fibres over P_2O_5.)

pattern at 66% r.h. (Plate IV of Part I) and an orthorhombic *C* pattern at 44% r.h. (Plate I(a)) as was expected. This fibre was then dried for 6 hours over P_2O_5; when re-examined it gave *hexagonal C* for all humidities from 32 to 75% r.h. (Plate III). Another fibre gave hexagonal *C* at 66% r.h., but after standing several days at 92% r.h. it gave orthorhombic *C* at 66% r.h. Thus the effect is not solely one of the water content of the fibre at the time of obtaining the pattern. Hysteresis effects are not removed by equilibrating for several days at constant humidity.

4. Measurement of Diffracted Intensity

A recording microdensitometer was used for intensity measurement, as described in Part I for crystalline diffraction patterns. However with semicrystalline diffraction patterns somewhat different problems are involved in the intensity measurements. First, since the intensity is observed as spots in the centre of the pattern and streaks in the outer part of the pattern, it is necessary to relate these two types of data.

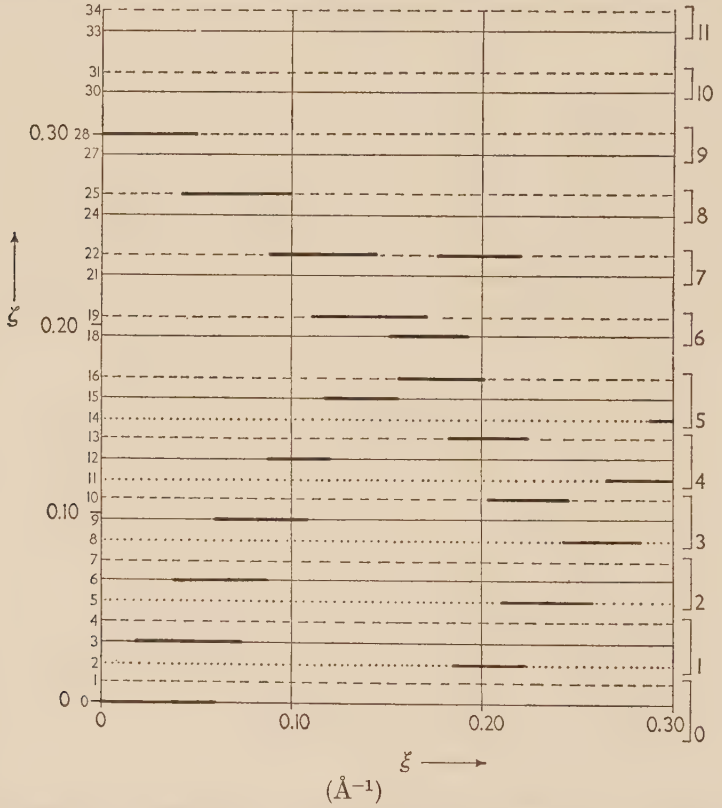


FIG. 2. Layer lines in the diffraction pattern of *C* DNA non-integral helix. On the left, numbering of layer lines for pitch 31·0 Å and internucleotide separation 3·32 Å is shown. Heavy lines superimposed on the layer lines indicate the regions of strong Bessel function contributions. The three major systems of Bessel functions contributing are $m = 0$ (solid), $m = 1$ (dashed) and $m = -1$ (dotted). On the right, the grouping of "split" layer lines and numbering of layer lines used in this work are shown.

Second, in much of the streaky region, neighbouring layer lines are not resolved; disorientation and other effects are such that the closely-spaced or "split" layer lines due to the non-integral helix (Fig. 2; Cochran, Crick, & Vand, 1952; Franklin & Klug, 1955) merge into continuous regions of intensity.

Streak intensity was corrected to the scale of the spots using the equation (Wilkins, 1961):

$$E_{sp} = \frac{nF^2(hkl)}{2\pi\xi N f_m^2 A} \cdot E_{st}$$

where

E_{sp} = Integrated intensity of spot

E_{st} = integrated intensity per Å^{-1} of streak

n = number of reciprocal lattice points overlapping on the fibre diagram and contributing to the spot

N = number of molecules in the unit cell (In the present case the term molecule refers to one turn of the DNA double helix.)

$F(hkl)$ = structure factor for the group of molecules in the unit cell

f_m = scattering factor for one molecule

ξ = distance of reciprocal lattice point from rotation axis, in Å^{-1}

A = area of c axis projection of the unit cell in Å^2.

A measure of the integrated intensity of spots was obtained by taking densitometer traces radially across a reflexion arc and tangentially along it. The integrated intensity was taken as the product of the areas under these two traces, divided by the peak height of the trace. Intensity per unit length along a layer line was determined by taking traces along a layer line and across it at certain points. The width of each transverse trace was defined as the area under the trace divided by the peak height. The intensity per unit length at any point of the layer line was then taken as the product of the width and height of the layer line trace at that point. In regions where neighbouring layer lines merged, the total area under a transverse trace was divided into regions corresponding to the layer lines of an integral ninefold helix. No attempt was made to resolve the splitting of the layer lines: the three major Bessel function contributions were measured together (Fig. 2).

Since layer lines were not always resolved on the diffraction patterns, guidance was needed in taking densitometer traces along a "layer line." A Bernal chart was prepared for the diffraction pattern, with layer lines and intervals of ξ marked on the chart (Fig. 3). An integral ninefold helix was assumed. The chart was reduced photographically to the size of the diffraction photograph, using the 3·03 Å calcite calibration ring to obtain precise correspondence. The reduced chart was then laid over the diffraction pattern during densitometry; traces along layer lines showed sharp peaks marking known intervals of ξ.

The largest source of error was in assignment of a background level of intensity: scatter from water and amorphous DNA varied in an unknown manner over the pattern, and the merging of neighbouring layer lines made it impossible to measure a background between layer lines. The background level was drawn in by eye for each layer line; the estimated maximum error was 50% where layer lines were ill-defined, and 20% where they were well-defined. A correction (about 5%) was made for overlapping of weak layer lines by strong ones on the 8th layer line. A correction was necessary on the 8th and 9th layer lines for broadening of peaks due to disorientation; for convenience, it was applied by broadening the peaks of the calculated transform.

Intensities from four different patterns are plotted in Fig. 4. Spots were measured on two patterns, one orthorhombic and one hexagonal. Streaks were measured on

three patterns, one orthorhombic and two hexagonal. The results agree within experimental error, and support the assumption that there is no significant difference in the molecular configuration when the molecules are packed in the two different ways.

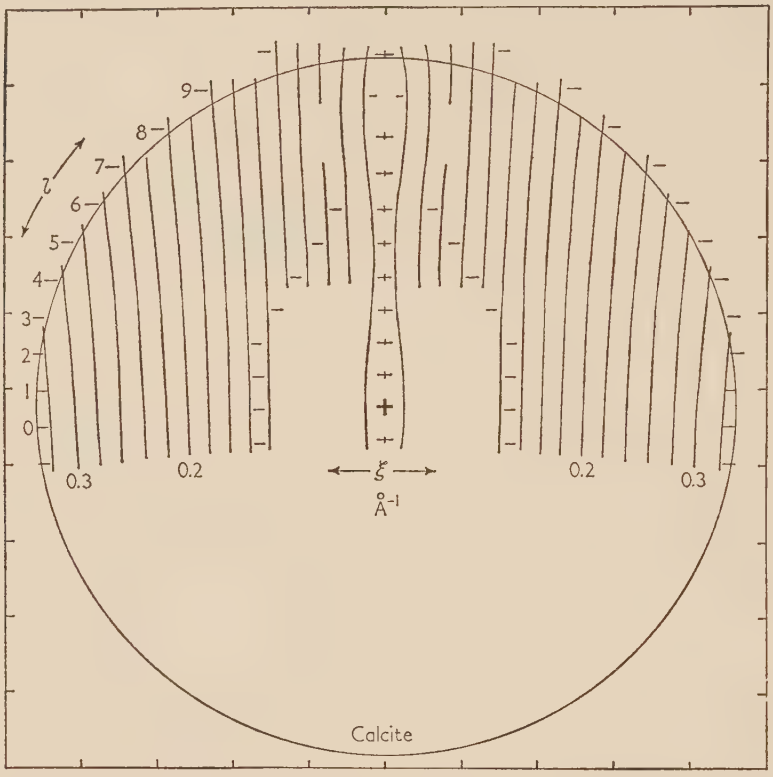


Fig. 3. Bernal chart used in densitometry of continuous transform regions on semicrystalline patterns. Short horizontal marks indicate the position of layer lines calculated for an integral ninefold helix of pitch 31·0 Å. Vertical lines mark intervals of ξ in Å^{-1} along the layer lines. The chart is designed for a flat film, CuKα radiation, and fibres tilted at 13° from perpendicular to the incident beam.

5. The Molecular Structure

(a) *Structure analysis*

The basic assumptions used in analysis of the *B* form (Part II) were used in analysis of the *C*. The DNA molecule was assumed to be a right-handed double helix, with a twofold rotation axis relating the phosphate-ester chains; the model was built in accordance with stereochemical data derived from the study of simpler compounds (Spencer, 1959*a*,*b*). In calculating the transform of the helix, scatter from water in the non-DNA regions was allowed for by an approximate theory employing scattering factors for atoms immersed in water. Analysis of the *C* form began with comparison of its diffraction pattern with that of *B*, for many similarities suggest that the two structures are closely related (Plate IV). The major differences are referred to below:

(1) The *C* configuration is a non-integral helix of slightly variable dimensions. Mean values of 3·32 Å for the nucleotide separation along the helix axis and 38·6° for the rotation between nucleotides were used in the model building.

(2) The *C* patterns show a region of strong intensity at $\xi = 0{\cdot}10$ Å^{-1} on the first layer line (X in Plate IV). This is a region of low intensity on *B* patterns. Consideration

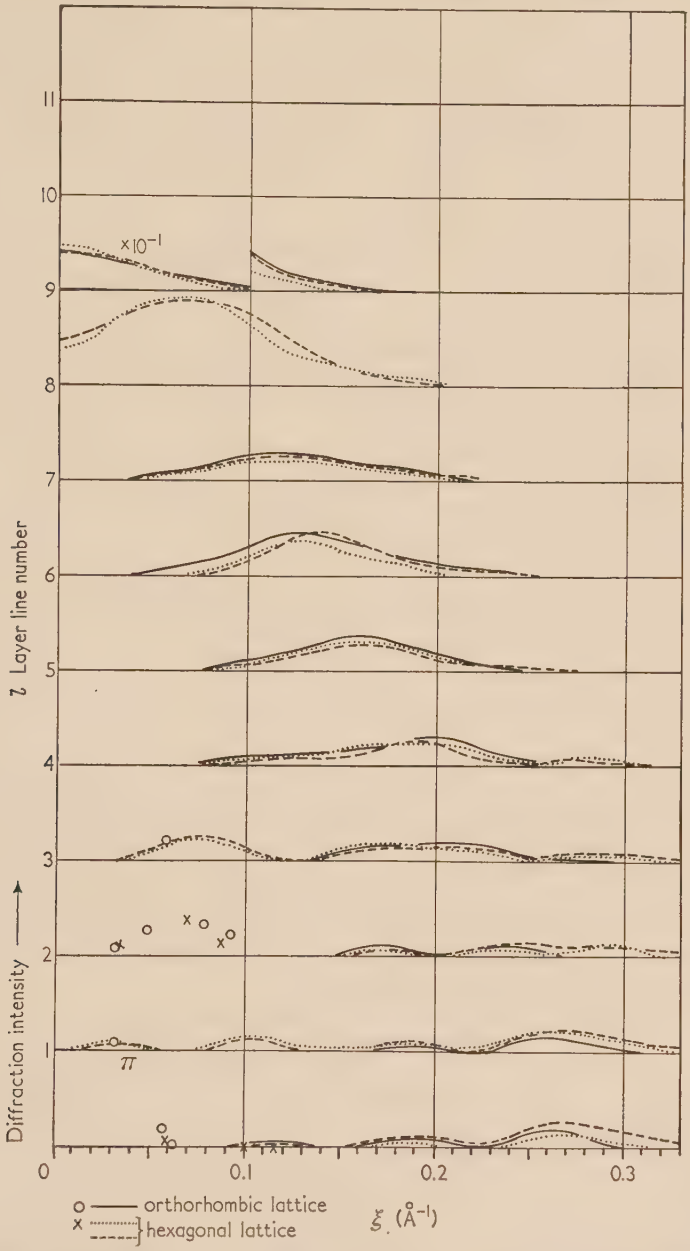


FIG. 4. Observed diffraction from hexagonal and orthorhombic patterns. Crosses and circles represent discrete reflexions; lines represent continuous transform regions of the diffraction. Results from different patterns were normalized so that the area under the trace along the meridional reflexion was the same in all cases.

of the Fourier transform of the *B* model (Part II, Fig. 3) shows that in this region the amplitude of scatter from the base interferes destructively with that from the phosphate-ester chain. The position of either of these groups in the helix could be

altered so as to cause the interference to become constructive. This would give increased intensity in this region, as observed on the *C* patterns. The distribution of intensity on the first four layer lines is otherwise similar on the *B* and *C* patterns: this suggested to us that the position of the phosphate-ester chain in the helix was the same in *C* as in *B*, because the sugar and phosphate groups diffract more strongly than the base in this region. Thus it is likely that the base position is different in the two configurations. Since only the central ($m = 0$) portion of the pattern is involved, the requirement that the position of the phosphate-ester chain is unaltered means that r and $\left(\frac{2\pi z}{c} - \phi\right)$ remain unchanged for each atom (Cochran *et al.*, 1952). In other words the position of the chain is unaltered when viewed as a smooth helix; although the position of the *atoms within the chain* may alter with respect to the coordinate system.

(3) The *C* pattern shows strong intensity on the layer line just below the strong meridional reflexion; the corresponding layer line is weak on the *B* pattern (Y in Plate IV). A radial projection (Klug *et al.*, 1958; Part II) indicated that this effect might be caused by tilting the bases 5° to 10° (around the dyad in the plane of the bases) from their position in the *B* model.

Thus qualitative considerations gave a rough picture of the nature of the *C* configuration. Trial calculations of transforms of bases indicated the direction in which the bases should be moved and tilted in order to give both effects (2) and (3). It was found stereochemically feasible to build such a model with the position of the phosphate-ester chain the same as in *B*, as required by (2). The stereochemistry and the fit of the calculated transform to the observed intensity were refined by detailed model building. Nine different molecular models were built and their transforms calculated. Model building confirmed that the phosphate-ester chain must be quite similar in position to that of *B*: any marked alteration of the phosphate-ester position gave a calculated transform that was incompatible with the observed diffraction.

The observed diffraction compared with that calculated from the final model is shown in Fig. 5. The main discrepancies are in the position of peaks on the 5th and 8th layer lines. The discrepancies are within the uncertainty of experimental measurements and the uncertainty of the assumptions made in building the model. As with the *B* model, the configuration derived here is probably accurate to within 0·5 Å in position of the heavily diffracting groups.

(b) *Discussion of the structure*

The base-pairs used in the model had the dimensions given by Spencer (1959*b*). The bases were moved along the dyad by 1·5 Å away from the helix axis relative to their positions in the *B* model; the 2° tilt of the bases for *B* was increased to 6° for *C*. The bases were twisted by 5° around a line through the pyrimidine C_4 atoms in order to relieve short interbase contacts and the short contact to the sugar C_2 atom. The base position and orientation are shown in Fig. 6. Two projections of a nucleotide of the final model are shown in Fig. 7. The covalent bond lengths and angles of the sugar-phosphate chain, and short van der Waals contacts between non-bonded atoms are shown on the diagram. The cylindrical coordinates of atoms in a residue are given in Table 2.

All bond lengths and angles are within the limits discussed in Part II. The increase in the deviation from linearity of the inter-base hydrogen bonds, due to the 5°

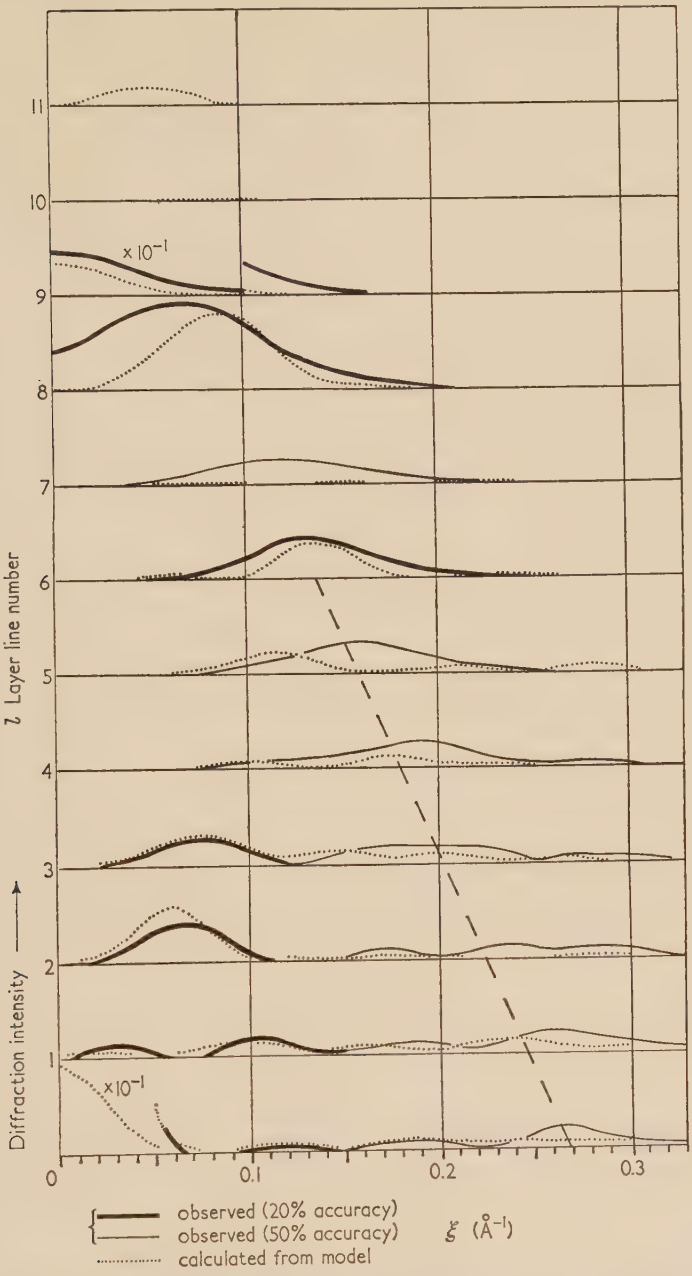


FIG. 5. Comparison of the observed diffraction (averaged from the data of Fig. 4) with that calculated from the final model. The observed was normalized to give the best overall correspondence with the calculated diffraction. Densitometer measurement of intensity on the 7th layer line was checked by visual estimation; it appears to be no more than one-third the intensity on the 6th layer line. The diagonal dotted line shows the position of the diagonal intensity ridge visible on the pattern (Plate I).

twisting of base-pairs, is small. For a 6° twist in a similar structure the maximum deviation has been calculated as 14° (Fuller, 1961). There are two somewhat short van der Waals contacts between bases in neighbouring nucleotide pairs but on different sides of the dyad. (These are not shown in the diagram.) One is a distance of 3·02 Å between the nitrogen atom of a guanine NH_2 group and an adenine C_2 atom; the other is 2·77 Å between the nitrogen atoms of two guanine NH_2 groups. The close contact of 2·97 Å between the sugar C_2 atom and the pyrimidine C_4 atom compares with 3·04 Å for the same contact in the *B* model; the slightly shorter distance would not significantly increase the repulsive forces involved. Other van der Waals contacts are within the limits defined in Part II. Present knowledge of bond angle distortion and close distances between atoms in van der Waals contact in large molecules is limited. While it may be said that the stereochemistry of the model does not conflict with data derived from structural studies of DNA components, it is possible that the model requires minor modification to relieve stereochemical strain.

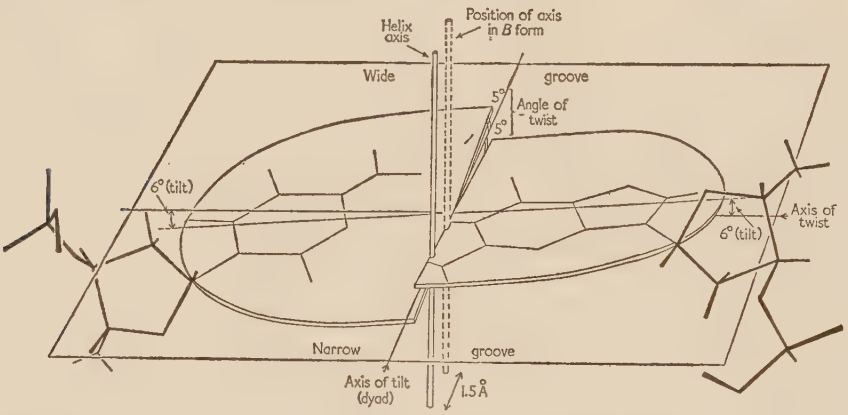


Fig. 6. Perspective drawing showing one nucleotide pair in the structure. The bases are shown as if drawn on solid plates which are rotated around the axes of tilt and twist by 6° and 5° respectively. The corresponding angles in the model for the *B* form are 2° tilt and 5° twist. The base-pair is moved 1·5 Å away from the helix axis relative to its position in the *B* form.

Comparison of Fig. 8 with Fig. 12 of Part II shows similarity between the *C* configuration and the *B* configuration of a single nucleotide regardless of its position in the helix. The sugar group has similar orientation in the two structures: the sugar ring is roughly perpendicular to the plane of the base, and the C_2 atom is puckered out of the plane of the other sugar atoms by 0·5 Å. The phosphate group also has a similar orientation in the two structures. Examination of three-dimensional models of the *B* and *C* structures confirms the similarity apparent in the Figures. As seen from the models, only slight rotations around bonds make the *B* and *C* nucleotides virtually identical.

It might be supposed that in converting from *B* to *C* the phosphate-ester chain would be stretched, for the molecule has been twisted about the helix axis so that the number of residues per turn is changed from 10 to $9\frac{1}{3}$, while the separation of nucleotides along the helix has decreased little. However, the effect of this twisting is counterbalanced by the movement of the bases away from the helix axis. The distance between successive C_1 atoms in the phosphate-ester chain is 4·87 Å in *B* and 4·93 Å in *C*.

Within the limits of the diffraction data the configuration of a single nucleotide appears to be the same for both the *B* and *C* structures. The only differences between

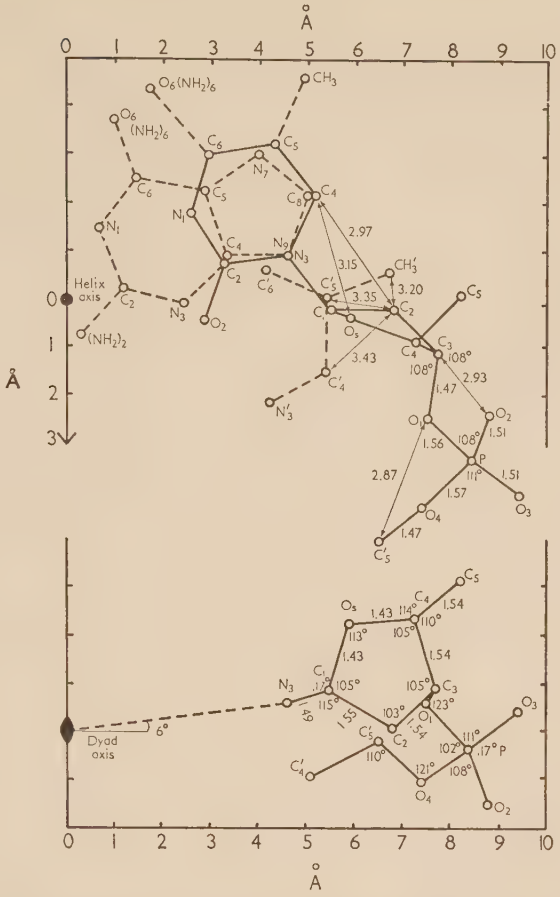


FIG. 7. Projections of a nucleotide of the *C* model. Top: projection down the helix axis; bottom: projection down the dyad axis. Covalent bond lengths and angles and short van der Waals contacts are shown.

FIG. 8. Part of the polynucleotide chain of the *C* model. The view is along the dyad axis of the upper nucleotide. Note the similarity of the nucleotide configuration to that of the *B* form (Fig. 12 of Part II).

TABLE 2

Coordinates of atoms in one nucleotide of the final C *model*

Group	Atom	r (Å)	ϕ (degrees)	z (Å)
phosphate	P	9·05	68·9	— 0·37
	O_1	7·91	71·8	0·60
	O_2	9·20	75·5	— 1·45
	O_3	10·32	67·3	0·40
	O_4	8·57	60·1	— 1·02
sugar	C_1	5·48	88·0	0·88
	C_2	6·79	88·1	0·05
	C_3	7·80	82·2	0·93
	C_4	7·28	83·4	2·37
	C_5	8·18	90·8	3·12
	O_5	5·90	86·6	2·24
purine	N_1	1·60	157·7	0·12
	C_2	1·12	103·3	0·28
	N_3	2·35	88·5	0·44
	C_4	3·34	106·1	0·45
	C_5	3·58	128·6	0·29
	C_6	2·90	150·2	0·12
	O_6, $(NH_2)_6$	3·86	164·7	— 0·03
	N_7	4·93	127·8	0·33
	C_8	5·38	114·6	0·51
	N_9	4·67	101·8	0·56
guanine	$(NH_2)_2$	0·83	17·2	0·28
pyrimidine	N_1	2·98	126·9	0·28
	C_2	3·30	103·0	0·46
	O_2	2·79	81·6	0·52
	N_3	4·67	101·8	0·56
	C_4	5·54	113·1	0·53
	C_5	5·34	127·6	0·35
	C_6	4·17	136·6	0·22
	O_6, $(NH_2)_6$	4·51	152·9	0·05
thymine	$(CH_3)_5$	6·74	133·9	0·29

For definition of r, ϕ, z see Fig. 1 of Part II. The helix is right-handed, with a translation of 3·32 Å and a rotation of 38·6 ° between successive nucleotides.

B and *C* are in the position and orientation of nucleotides within the helix. The model proposed for the *B* structure, at the time the *C* configuration was derived (model 2 of Part II), had a nucleotide configuration which was different from that of *C*. The final *B* model, with a reoriented phosphate group and a puckered sugar ring, was derived to a large extent independently of *C*. The similarity of these two independently derived *B* and *C* configurations is therefore further evidence that the proposed models are correct.

6. Molecular Packing

(a) *General discussion*

A helical macromolecule has two periodicities along its length: a shorter one, that of the residues along the helix, and a longer one, that of the helix pitch itself. If there is a well-defined ridge-and-groove structure on the molecule, as in tobacco mosaic virus or DNA, a major factor determining the interaction of closely packed parallel molecules will be the interlocking of the ridges and grooves of neighbouring molecules (Franklin & Klug, 1956; Klug & Franklin, 1958; Marvin, 1958). As a first approximation, the ridge-and-groove structure may be regarded as being without the fine detail given by the residues of the helix: the problem is then, how do smooth helices interlock? This approximation may not be valid if there is an integral number of residues in a pitch length, for then the pattern of contacts between residues on neighbouring molecules will repeat in every pitch length, and as a result the arrangement of residues

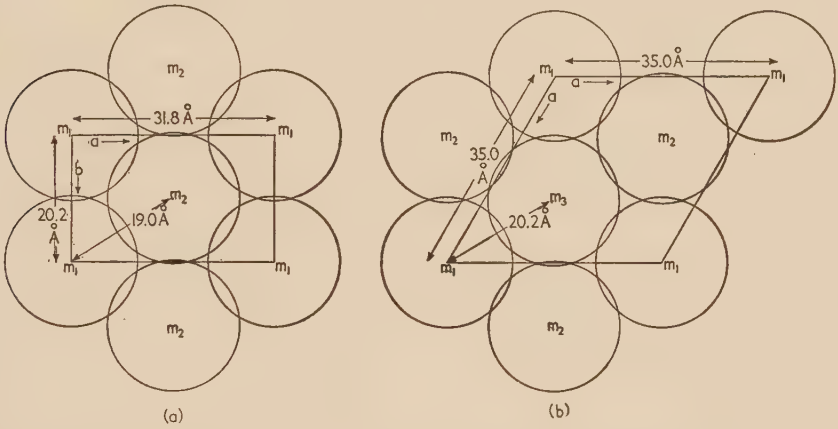


FIG. 9. Arrangement of the molecules in the unit cell: projection down the helix axis.
(a) orthorhombic packing.
(b) hexagonal packing.
Molecules m_1, m_2, and m_3 are displaced by different fractions of the helix pitch in the direction perpendicular to the diagram (see section 6). The lateral packing of molecules, without reference to their relative displacements along the helix axis, is not very different in the hexagonal and orthorhombic lattices. The circles have a radius equal to that of the furthest projecting atom, the phosphate O_3.

in the helix is likely to cause the molecules to take up specific orientations (Part II). However, when the number of residues per turn is not integral the pattern of contacts does not repeat every pitch length (Elliott & Malcolm, 1958), and as a result molecules are more likely to behave as interlocking smooth helices. Molecules which are packed as if they were smooth helices are said to have screw disorder (Klug & Franklin, 1958).

In studying the packing of DNA with screw disorder one may regard the molecule as a solid axial rod around which the two helical phosphate-ester ridges are wrapped; although the sequence of atoms is in opposite directions on the two chains, the two ridges may be considered equivalent for packing purposes. The position of each ridge may be defined by the helix passing through points midway between the O_2 and O_3

atoms of the phosphate groups. If P is the pitch of the helix, the two ridges are separated in the helix axis direction by distances xP and $(1 - x)P$; $(x \leqslant \frac{1}{2})$. Two parallel close-packed double helices have a relative displacement dP in the axis direction when the ridges and grooves interlock. In general there will be two interlocking positions: $d = (1 - x)/2$, and $d = 0$. If x is greater than $\frac{1}{3}$, the closer packing occurs when $d = (1 - x)/2$; if x is less than $\frac{1}{3}$, closer packing occurs when $d = 0$; if $x = \frac{1}{3}$,

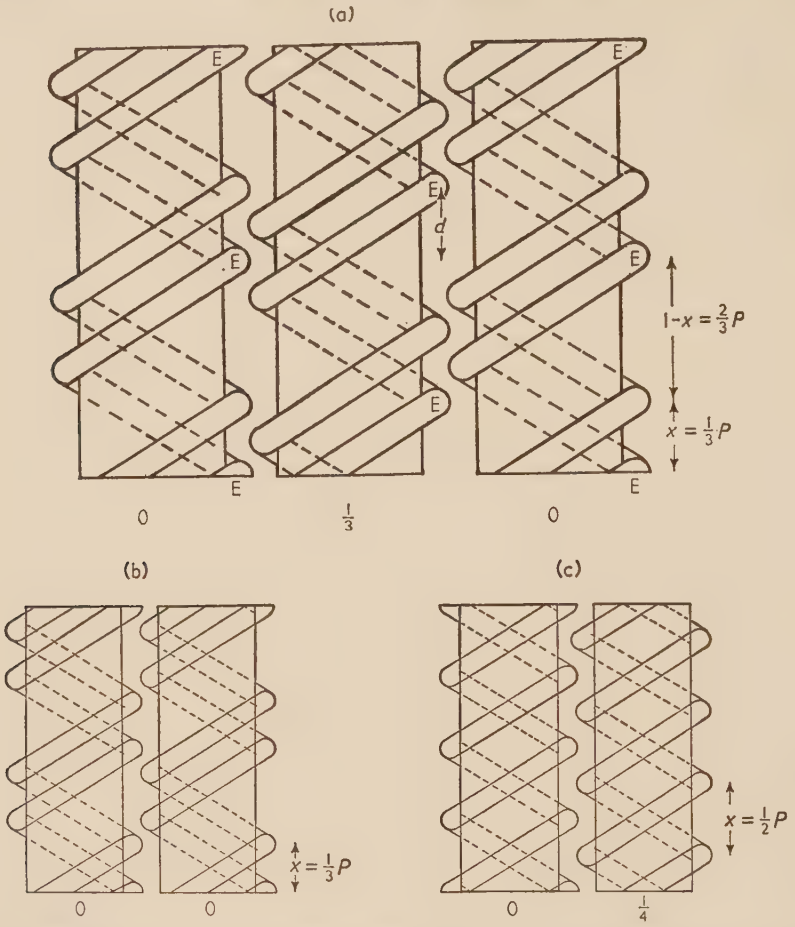


Fig. 10. Interlocking of smooth helices. A section along a line joining the helix axes is shown schematically. The letter E indicates equivalent points on the molecules, one pitch length apart.

(a) Group of two molecules with $x = \frac{1}{3}$. Molecules pack at heights 0, $\frac{1}{3}$, 0. This scheme is comparable to cubic close packing of equivalent spheres.

(b) Interlocking of two molecules with $x = \frac{1}{3}$ at the same height.

(c) Interlocking of two molecules with $x = \frac{1}{2}$. If the two ridges are considered to be equivalent, a random translation of any molecule by $\pm \frac{1}{2}$ would not disturb the packing.

equally close packing occurs for $d = \frac{1}{3}$ and $d = 0$. If $x = \frac{1}{2}$, there is no interlocking for $d = 0$; at $x = \frac{1}{2}$ either ridge may pack equally well into either groove, so a random axial displacement of $\frac{1}{2}$ is possible. Packing of smooth helices is illustrated in Fig. 10. For both the B and C configurations of DNA, $x = 0{\cdot}44$ and $d = 0{\cdot}28$.

One way of close-packing DNA molecules is in an orthorhombic lattice with two molecules m_1 and m_2 passing through the unit cell (Fig. 9(a)). The value of d for

molecules m_1 and m_2 can be adjusted to permit interlocking; the value of d for two m_1 molecules is 0, so they can also interlock. If x is greater than $\frac{1}{3}$, the m_1m_2 distance will be less than the close-packed m_1m_1 distance.

Another way of close-packing double helices is in a hexagonal lattice, with three molecules m_1, m_2, and m_3 passing through the unit cell (Fig. 9(b)). If $x = \frac{1}{3}$, the molecules can have heights 0, $\frac{1}{3}$, $\frac{2}{3}$, (Fig. 10(b)). $d = \frac{1}{3}$ for any two neighbours, and as a result the molecules can pack with equal distances between all neighbours. For the general case, the three molecules will have heights 0, h and $-h$ in the unit cell. Any two neighbours will have relative displacement of either h or $2h$. Since d gives the closest packing for any two molecules standing alone, the most stable situation for three molecules with $x > \frac{1}{4}$, all equally spaced from one another, will be when h and $2h$ are both equally far from d, or $h = \frac{2}{3}d$. The most stable situation with $x < \frac{1}{4}$ is $h = \frac{1}{3}$, assuming the packing forces are symmetrical (see Fig. 11).

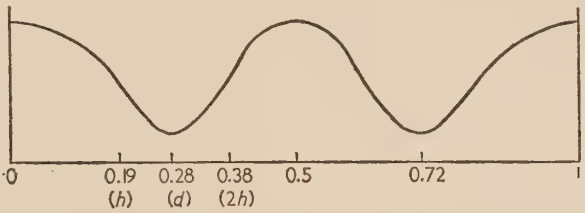


FIG. 11. Schematic representation of the closeness of packing for double helices. Abscissa: relative axial translation of the two molecules, in fraction of the pitch length. Ordinate: closeness of approach of the centres of the two molecules for a given relative translation; arbitrary unit. The curve shown is for a double helix with narrow groove width $x = 0{\cdot}44$. Two parallel helices pack closest at relative translation $d = 0{\cdot}28$ or $0{\cdot}72$, shown by the minima in the graph. Three parallel helices with their axes equidistant from each other cannot all have the relative translation d to each other unless $d = \frac{1}{3}$. In general they will have relative translations h or $2h$; both these will be equidistant from d or from $\frac{1}{2}$ for closest packing, assuming the curve is symmetrical.

(b) *Experimental data*

For *C* DNA, screw disorder is likely because the helix is non-integral. The presence of sharp spots in the centre of the pattern and streaks elsewhere is compatible with the presence of this disorder (section 2). Therefore intermolecular contacts were calculated assuming complete screw disorder.

For orthorhombic *C*, systematically absent intensities indicate that $d = 0{\cdot}25 \pm 0{\cdot}025$ (section 3); this is consistent with the value of 0·28 deduced from packing considerations. The closest contacts between interlocking molecules m_1 and m_2, 19·0 Å apart, is 2·9 Å between oxygen atoms. The closest contact between molecules m_1, 20·2 Å apart, is 2·5 Å. Local distortion of the molecules could readily increase the interatomic distances to a permissible value of 2·8 Å. Only a fraction of the unit cells will contain distortions.

For hexagonal *C*, positions and intensities of spots on the equator and 2nd layer line show that molecules m_1,m_2,m_3 (Fig. 9(b)) either have heights 0, $\frac{1}{6}$, $\frac{1}{3}$; or 0, $\frac{1}{3}$, $\frac{2}{3}$. The presence of streaks on the 1st and 3rd layer lines shows that the molecules are randomly displaced by $\frac{1}{2}P$ in the helix axis direction (section 3). This random displacement is possible because x is not far from $\frac{1}{2}$. If either of the molecular arrangements indicated by the spots is present, such a random displacement would give $h = \frac{1}{6}$ and $h = \frac{1}{3}$. Faint signs of spots on the 1st and 3rd layer lines show that part of the fibre consists of regular repeat units of molecules at approximate

heights 0, $\frac{1}{6}$, $-\frac{1}{6}$. This value of $\frac{1}{6}$ is consistent with the value of $h = \frac{2}{3}d = \cdot 19$ deduced from packing considerations. If screw disorder is present, centres of O_3 atoms on neighbouring molecules would sometimes be as close as 0·5 Å. Such short distances could be avoided in one unit cell by so orienting the molecules about their axes that phosphate groups interlock. If this is so the non-integral nature of the helix will in general cause butting of phosphate groups in neighbouring unit cells. Evidently a fairly considerable local distortion of the molecules is required where close contacts occur.

Calculation shows that if the molecules had heights 0, $\frac{1}{3}$, $\frac{2}{3}$, there would be no contacts less than 2·9 Å. It is not clear why this arrangement does not occur. We have assumed that the most likely arrangement of DNA molecules is that in which the molecules are closely packed and undistorted. This would be so if van der Waals forces were dominant in determining packing. In the case of DNA, however, electrostatic interaction will be considerable and may cause our simple approach to be invalid. The hexagonal packing arrangement for *C*, which gives rise to stereochemical difficulties, is the same as that found in the highly-hydrated form of *B* DNA (Feughelman *et al.*, 1955). While the separation of the molecules of *B* DNA is sufficient to avoid short inter-molecular distances, the question remains why the molecules do not form the apparently favourable arrangement 0, $\frac{1}{3}$, $\frac{2}{3}$. Presumably electrostatic interaction does not favour this arrangement. However, the double-helical DNA-like complex of polyribo-inosinic and cytidylic acids adopts this arragement (private communication from Dr. D. R. Davies) when it crystallises as a twelvefold helix.

We wish to thank Professor J. T. Randall, F.R.S., Drs. F. L. Horsfall, Jr. and J. Burchenal for encouragement; Drs. R. Langridge, H. R. Wilson, G. Zubay and Mr. W. Fuller for discussion; Mr. R. Larner for assistance with preparing DNA; Mr. P. J. Cooper for chemical analysis of DNA; I.B.M. United Kingdon, Ltd. and I.B.M. World Trade Corporation, and the University of London Computer Unit for provision of free computer time; Dr. R. Langridge and Mr. W. Fuller for providing computer programmes. One of us (D.A.M.) acknowledges award of a National Science Foundation Predoctoral Fellowship. The work at King's College was aided by a grant from The Rockefeller Foundation and that at Sloan-Kettering Research Institute by a grant from The Commonwealth Fund.

REFERENCES

Amoros, J. L., Canut, M. L. & De Acha, A. (1960). *Z Krist.* **114**, 39.
Cochran, W., Crick, F. H. C. & Vand, V. (1952). *Acta Cryst.* **5**, 581.
Davies, D. R. (1960). *Nature*, **186**, 1030.
Elliott, A. & Malcolm, B. R. (1958). *Proc. Roy. Soc.* A, **249**, 30.
Feughelman, M., Langridge, R., Seeds, W. E., Stokes, A. R., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F., Barclay, R. K. & Hamilton, L. D. (1955). *Nature*, **175**, 834.
Franklin, R. E. & Gosling, R. G. (1953*a*). *Acta Cryst.* **6**, 673.
Franklin, R. E. & Gosling, R. G. (1953*b*). *Acta Cryst.* **6**, 678.
Franklin, R. E. & Gosling, R. G. (1955). *Acta Cryst.* **8**, 151.
Franklin, R. E. & Klug, A. (1955). *Acta Cryst.* **8**, 777.
Franklin, R. E. & Klug, A. (1956). *Biochim. biophys. Acta*, **19**, 403.
Fuller, W. (1961). *J. Mol. Biol.* **3**, 175.
Hosemann, R. (1950*a*). *Z. Phys.* **128**, 1.
Hosemann, R. (1950*b*). *Z. Phys.* **128**, 465.
Hosemann, R. (1951). *Acta Cryst.* **4**, 520.
James, R. W. (1954). *The Optical Principles of the Diffraction of X-rays*. London: Bell.

Klug, A., Crick, F. H. C. & Wyckoff, H. W. (1958). *Acta Cryst.* **11**, 199.
Klug, A. & Franklin, R. E. (1958). *Disc. Faraday Soc.* **25**, 104.
Langridge, R., Seeds, W. E., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F. & Hamilton, L. D. (1957). *J. Biophys. Biochem. Cytol.* **3**, 767.
Langridge, R., Seeds, W. E., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F. & Hamilton, L. D. (1960*a*). *J. Mol. Biol.* **2**, 19.
Langridge, R., Marvin, D. A., Seeds, W. E., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F. & Hamilton, L. D. (1960*b*). *J. Mol. Biol.* **2**, 38.
Marvin, D. A. (1958). *Disc. Faraday Soc.* **25**, 203.
Marvin, D. A. (1960). Ph.D. Thesis, University of London.
Marvin, D. A., Spencer, M., Wilkins, M. H. F. & Hamilton, L. D. (1958). *Nature*, **182**, 387.
Pregl, F. (1951). *Quantitative Organic Microanalysis*. London: Churchill.
Spencer, M. (1959*a*). *Acta Cryst.* **12**, 59.
Spencer, M. (1959*b*). *Acta Cryst.* **12**, 66.
Stokes, A. R. (1955). *Progress in Biophysics*, **5**, 140.
Stokes, A. R. & Wilson, A. J. C. (1944). *Proc. Phys. Soc.* **56**, 174.
Wilkins, M. H. F. (1961). In preparation.
Wilkins, M. H. F., Stokes, A. R. & Wilson, H. R. (1953). *Nature*, **171**, 737.
Wyckoff, H. W. (1955). Ph.D. Thesis, Massachusetts Institute of Technology.

J. Mol. Biol. (1961) **3**, 566–584

La Structure du Poly-L-γ-Glutamate de Benzyle en Solution. Configuration en Hélice Différente de l'Hélice α et Transitions entre Formes Hélicoïdales †

V. Luzzati, M. Cesari,‡ G. Spach, F. Masson et J. M. Vincent

Centre de Recherches sur les Macromolécules, 6 rue Boussingault, Strasbourg, France

(*Received 29 March 1961*)

The structure of poly-γ-benzyl-L-glutamate (PLBG) in solution in dimethylformamide, pyridine and *m*-cresol has been studied by X-ray diffraction techniques.

Dilute isotropic solutions have been investigated by small-angle X-ray scattering experiments performed on an absolute scale. These experiments confirm that the molecules have a rod-like configuration but show that the mass per unit length of the rods is 30% smaller than in the α-helix.

This observation is confirmed by the study of the liquid-crystalline *cholesteric phase* that PLBG forms in solution in *m*-cresol and pyridine at fairly high concentrations. The mass per unit length of the rods in *m*-cresol is 30% smaller than for the α-helix, as in the dilute solutions, while in pyridine it closely agrees with that of the α-helix.

At even higher concentrations two more phases are observed. The *paracrystalline phase* that exists in *m*-cresol is a hexagonal array of rigid rods: the mass per unit length of the rods is that of the α-helix. *A complex phase* is observed in dimethylformamide and pyridine whose structure is shown to be formed by bundles of three α-helices, twisted into a three-stranded secondary helix, providing an example of the "coiled-coil" described by Crick.

These observations indicate that the structure of PLBG is α-helical in concentrated solutions, but that in dilute solutions the structure is of a different kind: the experimental data for dilute solution are consistent with the 3_{10}-helix, which has been shown to be almost as stable as the α-helix.

This helical polymorphism is quite unexpected since all the previously described hydrodynamic, light-scattering and optical rotation experiments have been interpreted as showing that the structure of PLBG is that of the α-helix in dilute solution (in the solvents we have used).

These experiments are discussed and it is shown that the light scattering observations, when correctly interpreted, are in better agreement with the 3_{10}-helix than with the α-helix: the hydrodynamic experiments cannot distinguish between these two structures.

† *Note ajoutée aux épreuves.* J. Marchal et E. Marchal ont repris récemment l'étude des spectres d'absorption diélectrique des polypeptides en solution (Colloque Ampere, 1960, *Arch. Sci., Genève*, fasc. spéc., p82). Ils ont montré que le comportement du PLGB en solution dans le chloroforme est celui d'un bâtonnet flexible, dont $h > 1{,}5$ Å (Colloque Ampere, 1961, *Arch. Sci., Genève*, sous presse).

‡ Boursier de Olin Mathieson Chemical Corp.

1. Introduction

On sait depuis les travaux classiques de Astbury (voir Low, 1953), que, dans une importante famille de protéines fibreuses, les chaînes polypeptidiques peuvent adopter deux configurations, l'une appelée α, où les molécules sont maintenues repliées sur elles-mêmes par un ensemble de liaisons hydrogène intra-moléculaire, l'autre, dite β, où les chaînes polypeptidiques sont au contraire étirées.

Pauling, Corey & Branson (1951) ont généralisé cette notion de configuration α, en proposant un ensemble de structures hélicoïdales (voir également Bamford, Elliott & Hanby, 1956), modèles précis accessibles aux vérifications expérimentales. Ces dernières n'ont pas été longues à venir: dès 1951 Perutz, suivi par d'autres (Bamford *et al.*, 1956), a reconnu dans les diagrammes de diffraction des rayons X du poly-L-glutamate de benzyle et de quelques protéines, certaines des caractéristiques de ces structures hélicoïdales. En fait une de ces structures, appelée par Pauling & Corey "hélice α" a pris, dès le début, une place prépondérante.

Dans ce succès de l'hélice α le poly-L-γ-glutamate de benzyle (abrégé PLGB dans la suite) a joué un rôle privilégié. Utilisées par Perutz dans la première vérification cristallographique du modèle, les fibres de PLGB ont été étudiées de manière plus précise par la suite: la structure s'est avérée être en excellent accord avec celle de l'hélice α de Pauling & Corey. C'est également avec des échantillons orientés de ce polypeptide que Ambrose & Elliott (1951) ont mis en évidence le dichroïsme dans l'infra-rouge, et ont montré que les liaisons hydrogène entre C=O et N—H sont sensiblement parallèles à l'axe de la fibre: cette orientation est caractéristique des structures hélicoïdales (voir Bamford *et al.*, 1956).

La structure du PLGB a été étudiée également en solution diluée par des mesures d'absorption dans l'infra-rouge (Ambrose & Elliott, 1951), de pouvoir optique rotatoire (Yang & Doty, 1957; Downie, Elliott, Hanby & Malcolm, 1957), de viscosité (Doty, Bradbury & Holtzer, 1956; Yang, 1958*a*; Spach, 1960), de diffusion de la lumière (Doty *et al.*, 1956; Spach, 1960), de biréfringence d'écoulement (Yang, 1958*b*), de moment diélectrique (Wada, 1960), de diffusion de translation (Spach, 1960), de diffusion de rotation (Tinoco, 1957; O'Konski, Yoshioda & Orttung, 1959; Marchal & Marchal, 1959; Benoit & Boeckel, en préparation).

Il a été ainsi montré que le PLGB peut adopter deux configurations, selon la nature du solvant: l'une, du type α, existe dans certains solvants modérément polaires (diméthylformamide, pyridine, *méta*-crésol), l'autre, lâche et chaotique, du type "chaîne statistique", existe au contraire dans certains solvants très polaires (acide dichloracétique, hydrazine). Les propriétés de la configuration du type α indiquent clairement que les molécules ont la forme de bâtonnets rigides et que la chaîne polypeptidique est repliée de manière régulière, vraisemblablement hélicoïdale. La tentation de reconnaître dans cette structure l'hélice α de Pauling & Corey était grande: l'accord raisonnable de certaines mesures hydrodynamiques et de diffusion de la lumière avec les dimensions caractéristiques de l'hélice α a semblé justifier la conclusion, adoptée unanimement par la suite, que la structure du PLGB est, en solution dans ces solvants, celle de l'hélice α (Doty, 1959).

Nous pouvons avouer qu'en entreprenant l'étude des solutions diluées et isotropes de PLGB nous cherchions surtout à étalonner la technique de la diffusion centrale des rayons X, mesurée à l'échelle absolue, mise au point de fraîche date (Luzzati, 1960): nous étions certains, sur la foi des travaux précédents, d'avoir affaire à des hélices α.

Dès le début, toutefois, la valeur de la masse linéaire des bâtonnets, mesurée par notre technique, s'est obstinée à rester considérablement inférieure à celle de l'hélice α. Ce résultat, accueilli avec bien des réticences, a dû finalement être accepté, lorsque l'expérience acquise avec d'autres substances (Luzzati, Nicolaïeff & Masson, 1961; Luzzati, Witz & Nicolaïeff, 1961*a*, *b*) a prouvé que l'écart dépasse largement l'erreur des mesures.

Entretemps, intéressés par les phénomènes optiques remarquables que Robinson a observés (1956; Robinson, Ward & Beevers, 1958) dans les phases liquide-cristallines de certains systèmes PLGB-solvant, nous avions commencé à examiner les solutions concentrées de PLGB. Dans ces solutions également les observations expérimentales ont mis en évidence un ensemble de phénomènes bien plus complexes que ceux que nous nous attendions à trouver.

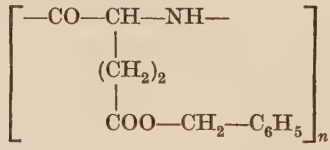
Guidés par l'expérience acquise dans l'étude des colloïdes d'association (Luzzati, Mustacchi, Skoulios & Husson, 1960; Husson, Mustacchi & Luzzati, 1960), dans les nucléoprotéines (Luzzati & Nicolaïeff, 1959) et dans l'acide désoxyribonucléique (Luzzati, Nicolaïeff & Masson, 1961), nous avons alors décidé d'effectuer une étude systématique, en fonction de la concentration, de la structure des phases que forme le PLGB avec ses solvants. Nous décrivons ici des résultats obtenus avec la diméthylformamide, la pyridine et le *méta*-crésol.

2. Technique Experimentale

(a) *Le polypeptide*

Les différents échantillons de PLGB ont été préparés selon le mode opératoire décrit par Blout & Karlson (1956). Les polymérisations ont été effectuées en solution dans le dioxanne et en utilisant comme initiateur le méthanolate de Na.

La formule chimique du polymère est



la masse moléculaire est n (219,16), le nombre d'électrons n (116).

La masse moléculaire des différents échantillons a été déterminée par diffusion de la lumière, et, dans quelques cas, par des mesures de viscosité, préalablement étalonnées par diffusion de la lumière (Spach, 1960).

Ayant à décrire, dans la suite, des structures hélicoïdales, nous les caractériserons par la distance h entre deux monomères successifs, projetée sur l'axe de l'hélice: pour l'hélice α $h = 1{,}5$ Å.

Les solvants ont été soigneusement déshydratés, puis purifiés par plusieurs distillations.

(b) *Etude des solutions diluées et isotropes*

Les solutions ont été préparées en mélangeant des quantités déterminées de polypeptide et de solvant. Il a été parfois assez délicat d'obtenir des mélanges homogènes: avec les solutions les plus visqueuses on a eu recours à une agitation mécanique, et à une élévation de la température. Le transvasement des solutions dans le porte-échantillon de l'appareil de rayons X a été effectué avec l'aide d'un dispostif qui rend négligeables les pertes du solvant dues à l'évaporation. Toutes les expériences ont été effectuées à une température voisine de 25°C.

Les solutions ont été étudiées au moyen de la diffusion centrale des rayons X, mesurée à une échelle absolue, selon la technique que nous avons décrite ailleurs (Luzzati, 1960; Luzzati, Nicolaïeff & Masson, 1961).

La notation est ici celle des mémoires précédents:

s est le paramètre qui mesure les angles de diffraction: $s = \frac{2 \sin \theta}{\lambda}$, exprimé en $Å^{-1}$.

c (ou c_e) est la concentration, mesurée par le rapport entre la masse (ou le nombre d'électrons) du PLGB et la masse (ou le nombre d'électrons) de la solution.

δ (ou ρ_0) est la densité (ou la densité électronique) du solvant exprimée en g cm^{-3} (ou él. $Å^{-3}$).

$\bar{v}$ (ou ψ) est le volume spécifique partiel du PLGB dans le solvant: $\bar{v}$ est exprimé en g^{-1} cm^3, ψ en $él.^{-1}$ $Å^3$.

$\frac{m}{l}$ (ou μ) est la masse linéaire du PLGB, mesurée par la masse en grammes (ou le nombre d'électrons) d'un segment de bâtonnet de longueur unitaire:

$$\mu \text{ (él. Å}^{-1}) = \frac{m}{l} (\text{g cm}^{-1}) \times 3{,}19 \times 10^{15} = 116/h$$

R_c est le rayon de giration, autour de l'axe du bâtonnet, de la différence entre la densité électronique à l'intérieur du bâtonnet et la densité électronique du solvant: R_c est exprimé en Å.

Rappelons brièvement que le résultat de ces expériences se présente, après normation, sous la forme d'une fonction $j_n(s)$, indépendante des conditions expérimentales, dont on connaît l'expression en fonction de la structure de l'échantillon. Lorsque ce dernier est une solution isotrope de bâtonnets très longs et rigides, tous de même section, distribués sans corrélations en position et en orientation, l'expression de $j_n(s)$, près de l'origine ne dépend que de la masse linéaire des bâtonnets et du rayon de giration, autour de l'axe:

$$j_n(s) = A \tfrac{1}{2} \exp(-\pi^2 R_c^2 s^2) K_0(\pi^2 R_c^2 s^2) \quad (1)$$

$$A = c_e \mu (1 - \rho_0 \psi)^2 \quad (2)$$

(K_0 est une fonction de Bessel d'argument imaginaire, voir Watson, 1952).

L'analyse des courbes expérimentales est facilitée par l'emploi d'une représentation doublement logarithmique. Chaque courbe expérimentale log $j_n(s)$, tracée en fonction de log s, est déplacée, sans rotations, par rapport à la fonction théorique (voir 1), log $[\frac{1}{2} \exp(-\pi^2 s^2) K_0(\pi^2 s^2)]$ tracée également en fonction de log s. Lorsqu'on réussit ainsi à faire coïncider les deux courbes, du moins pour s petit (et compte tenu des effets des corrélations), on trouve satisfaite une condition nécessaire pour que les particules de soluté aient la forme de bâtonnets: les déplacements d'abscisses et d'ordonnées mesurent alors R_c et A. Connaissant c_e, ψ et ρ_0, on peut déterminer la valeur de μ (2).

Le domaine de validité des équations (1) et (2) a été discuté ailleurs (Luzzati, 1960). On peut éliminer les effets des corrélations, qui en toute rigueur ne deviennent négligeables qu'à dilution infinie, en effectuant plusieurs expériences, à différentes concentrations, et en extrapolant les résultats à $c_e = 0$. Ce problème de l'extrapolation, ainsi que certaines modalités d'emploi des équations (1) et (2), ont été discutés récemment, à l'occasion de l'étude de l'acide désoxyribonucléique (Luzzati, Nicolaïeff & Masson, 1961).

(c) *Etude des solutions concentrées: phases mésomorphes et cristallines*

Ces solutions ont été préparées par mélange de quantités connues de polypeptide et de solvant: elles ont été examinées lorsque l'équilibre de concentration était atteint. Les concentrations ont été vérifiées par détermination de poids sec des échantillons après les expériences de rayons X.

Les diagrammes de diffraction des rayons X ont été obtenus avec des chambres à focalisation avec rayonnement monochromatique, à détection photographique, opérant sous vide (Luzzati & Nicolaïeff, 1959).

Les clichés contiennent, dans leur région centrale, quelques raies fines, ou parfois des bandes assez floues. Lorsque le réseau est hexagonal à deux dimensions, l'espacement ($Å^{-1}$) de la première raie s_1 varie avec la concentration c selon la loi (Luzzati, Nicolaïeff & Masson, 1961):

$$s_1^{-2} \times 10^{-16} = \frac{\sqrt{3}}{2} \frac{m}{l} \frac{1}{\delta} \left[c^{-1} - 1 + \delta \bar{v} \right] \qquad (3)$$

Connaissant $\bar{v}$ et δ, l'analyse de la variation de s_1 avec c permet de déterminer m/l.

Nous avons examiné également ces échantillons au microscope polarisant, en utilisant aussi bien les cuves employées pour les expériences de rayons X que des préparations enfermées entre lame et lamelle.

Toutes les expériences ont été effectuées à température ambiante (20°C environ), sauf dans quelques cas où la température est expressément indiquée.

3. Resultats

(a) *Volume spécifique*

La valeur de la masse linéaire, qu'on obtient par l'interprétation des expériences de diffusion centrale absolue, dépend de manière critique de la précision avec laquelle est connu le volume spécifique (eq. 2).

Nous avons mesuré le volume spécifique du PLGB dans les différents solvants en nous servant d'un picnomètre en pyrex, formé d'un bulbe auquel est soudé un tube capillaire gradué; le volume du picnomètre est de 10 cm^3 environ.

Les résultats des mesures sont portés dans le Tableau 1. La valeur moyenne a été calculée en pondérant les valeurs expérimentales par la concentration: la dispersion des points ne dépasse pas, en moyenne, 0,003 $cm^3\ g^{-1}$. Par ailleurs, Mitchell, Woodward & Doty (1957) ont obtenu, dans la diméthylformamide, $\bar{v} = 0{,}791 \pm 0{,}006$ à 25°C, valeur qui est en bon accord avec nos mesures.

Tableau 1

Volume spécifique du poly-L-glutamate de benzyle dans les différents solvants

Solvant	t °C	c %	$\bar{v}$ cm^3g^{-1}	$\overline{(\bar{v})}$ † cm^3g^{-1}
diméthylformamide	20,0	5,163	0,784	
		2,258	0,791	
		1,091	0,788	
		0,949	0,785	0·786
pyridine	25,1	3,874	0,788	
		2,857	0,787	
		1,554	0,786	0,787
méta-crésol	25,1	3,452	0,786	
		1,956	0,787	
		1,084	0,793	0,788

† valeur moyenne pondérée par la concentration.

(b) *Solution isotrope*

Cette phase existe dans les trois solvants aux concentrations les plus faibles. Au microscope polarisant, et au repos, elle est isotrope.

Les diagrammes de diffraction des rayons X contiennent, dans la partie centrale, une tache diffuse, que nous avons explorée au moyen de la technique décrite ci-dessus. Les résultats expérimentaux sont portés dans la Figure 1, dans la représentation

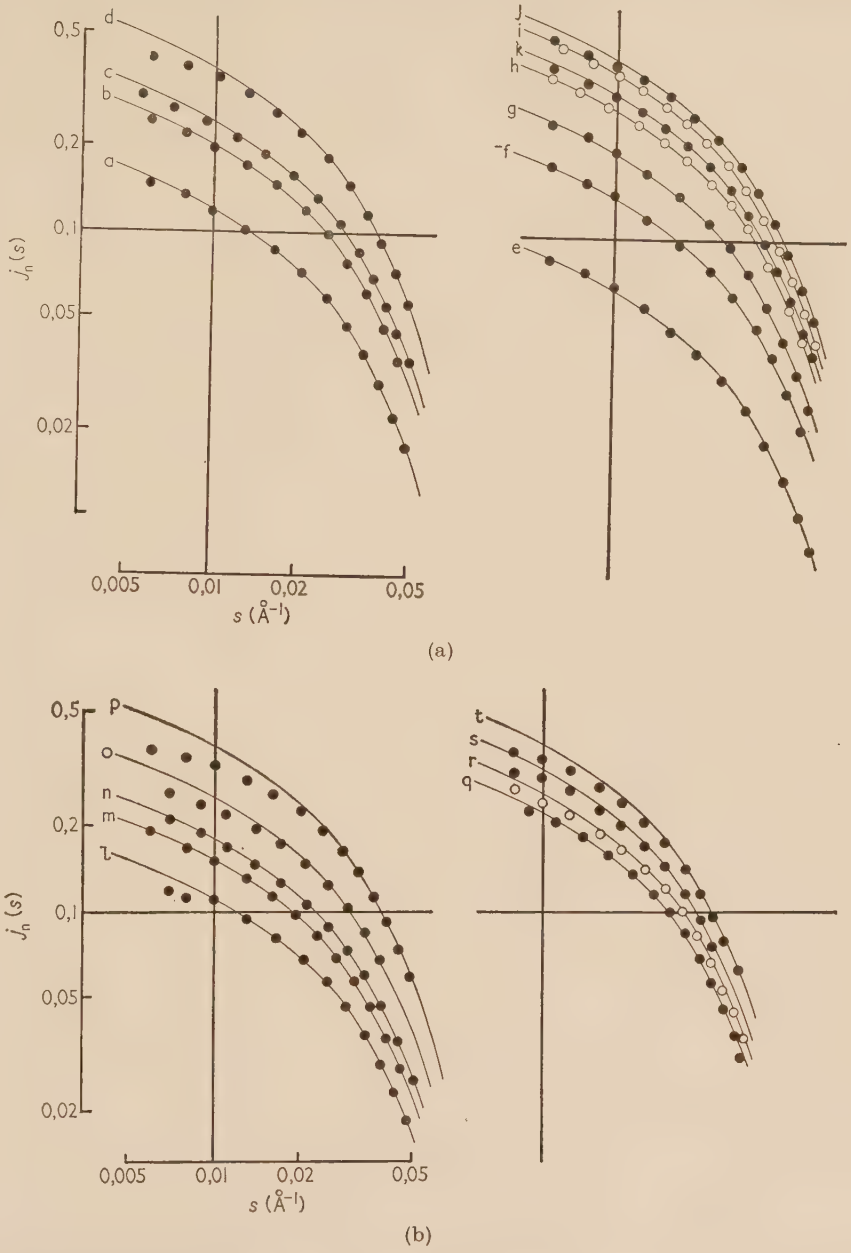


FIG. 1. Résultats expérimentaux des expériences de diffusion centrale à l'échelle absolue: les concentrations sont données dans le Tableau 2.

(a) diméthylformamide (b) pyridine (c) *méta*-crésol (*v.infra*)

doublement logarithmique habituelle. La forme des courbes est typique d'une solution de bâtonnets: nous avons tracé dans la Figure 1 la courbe théorique (1) qui

coïncide le mieux avec chaque courbe expérimentale, compte tenu des effets des corrélations. L'accord entre courbes théoriques et expérimentales est excellent: les valeurs de A/c_e sont indépendantes de la concentration. Dans chaque famille d'expériences l'écart entre $j_n(s)$ et la fonction théorique augmente avec la concentration, et s'accentue au voisinage de l'origine, ce que prévoit la théorie (Luzzati, Nicolaïeff & Masson, 1961). Les effets des corrélations dépendent en outre de la nature du solvant et de la masse moléculaire du polypeptide.

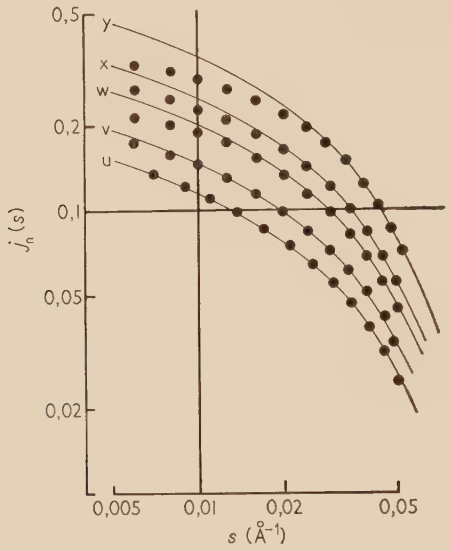


Fig. 1 (c)

Connaissant A/c_e, on peut calculer μ et h: nous avons effectué ce calcul pour chaque série d'expériences, en utilisant la valeur moyenne de A/c_e (Tableau 2). Afin de faciliter la discussion des erreurs, nous avons tracé h en fonction de $\bar{v}$, dans la Figure 2; nous y avons indiqué également la valeur la plus probable de h, correspondant à $(\overline{\bar{v}})$ du Tableau 1.

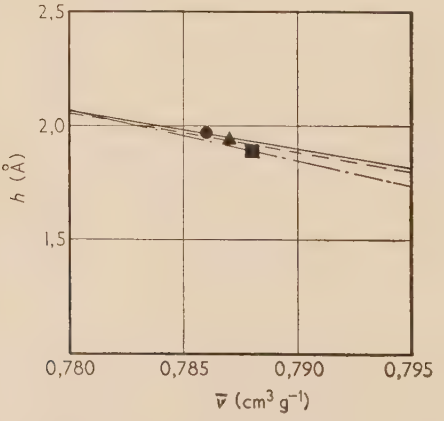


Fig. 2. Représentation de h, en fonction de $\bar{v}$ (voir Tableau 2). On a indiqué également la valeur la plus probable de h, correspondant à $(\overline{\bar{v}})$ du Tableau 1.

——●—— diméthylformamide ——▲—— pyridine ——■—— *méta*-crésol

TABLEAU 2

Résultats expérimentaux relatifs aux solutions diluées

	Solvant	ρ_0 él. Å^{-3}	M 10^3	c%	c_e%	R_c Å	$\overline{A}$	A/c_e	$\overline{(A/c_e)}$
a	diméthyl-	0,3108	35	1,86	1,80	4,6	0,060	3,33	
b	formamide			3,37	3,25		0,103	3,17	
c				4,10	3,96		0,123	3,11	
d				6,05	5,84		0,192	3,29	3,225
e			330	1,00	0,96	4,6	0,033	3,44	
f				2,05	1,98		0,070	3,53	
g				3,00	2,90		0,101	3,48	
h				4,00	3,86		0,143	3,70	
k				5,00	4,82		0,163	3,38	
i				6,00	5,79		0,192	3,32	
j				7,00	6,76		0,217	3,21	3,437
l	pyridine	0,3116	35	1,76	1,75	4,4	0,056	3,20	
m				2,28	2,27		0,076	3,35	
n				2,79	2,77		0,089	3,21	
o				4,00	3,98		0,125	3,14	
p				6,22	6,19		0,189	3,05	3,190
q			250	3,30	3,28	4,4	0,111	3,38	
r				3,95	3,93		0,130	3,31	
s				4,88	4,85		0,156	3,22	
t				5,88	5,85		0,192	3,28	3,297
u	*m*-crésol	0,3317	35	2,5	2,46	3,5	0,051	2,07	
v				3,5	3,45		0,065	1,88	
w				4,4	4,34		0,089	2,05	
x				5,9	5,82		0,111	1,91	
y				7,9	7,79		0,156	2,00	1,982

Il convient d'analyser la précision de ces résultats. Nous avons discuté ailleurs (Luzzati, Nicolaïeff & Masson, 1961) les erreurs inhérentes à la technique expérimentale: elles sont certainement inférieures à 5%. Nous avons discuté également de l'incertitude du choix de R_c, et de ses conséquences sur A/c_e: l'erreur sur R_c ne dépasse pas ici 0,15 Å, ce qui conduit, dans l'hypothèse la plus défavorable, à une erreur relative sur A/c_e voisine de 5%.

L'erreur due à $\bar{v}$ n'est pas grande: même en admettant que l'erreur commise sur $\bar{v}$ atteint 0,003 cm^3 g^{-1} (voir Tableau 1), il en résulte une imprécision dans la mesure de h qui n'est pas supérieure à 0,05 Å dans le cas de la diméthylformamide et de la pyridine, et à 0,07 Å dans le cas du *méta*-crésol.

Il reste à expliquer la petite variation de A/c_e avec la masse moléculaire du polypeptide (Tableau 2), et à faire un choix entre les différentes valeurs. Nous avons discuté de ce problème dans l'Appendice 1: nous avons montré qu'il convient de préférer les valeurs correspondant aux petites masses. Il s'agit, toutefois, de différences mineures qui, en première approximation, pourraient être négligées.

En conclusion *la valeur de h, dans les trois solvants, est voisine de 1,95 Å*; la différence avec la valeur correspondant à l'hélice α (h = 1,5 Å) dépasse largement l'erreur expérimentale.

(c) *Phase cholestérique*

Cette phase est d'aspect trouble et présente, même à l'oeil nu, de remarquables effets d'interférence. Au microscope polarisant on observe les textures que Robinson (1956; Robinson *et al.*, 1958) a décrites.

Les diagrammes de diffraction des rayons X contiennent, aux environs de $s = 4 \times 10^{-2}$ Å^{-1}, une bande floue qui se déplace vers les grands angles, et s'affine,

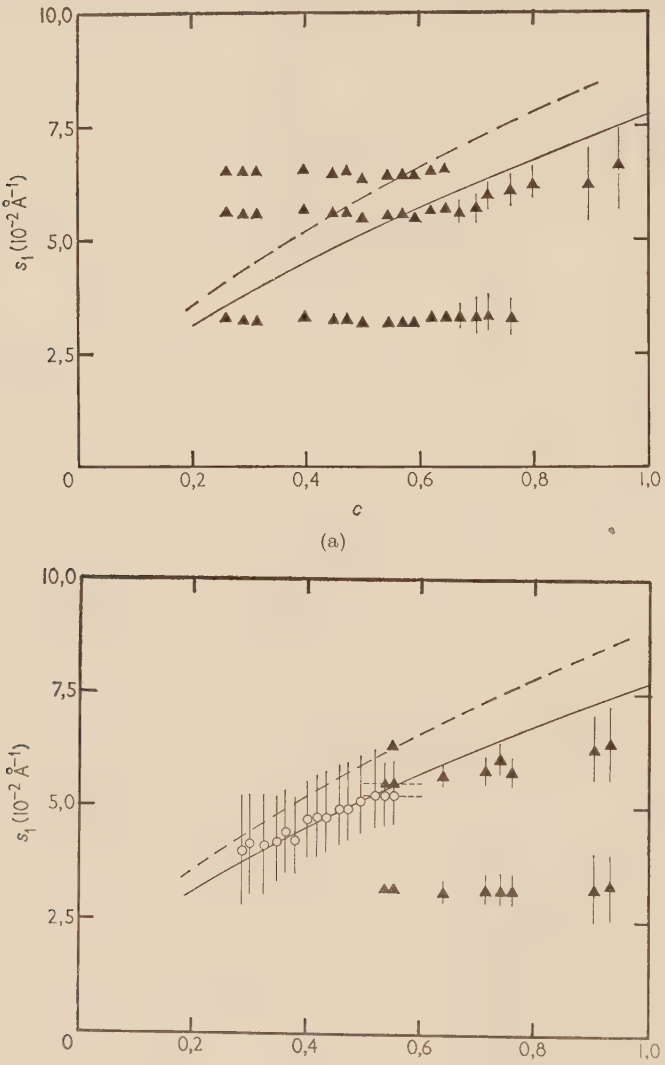


Fig. 3. Position des raies des phases organisées, en fonction de la concentration. Lorsqu'il s'agit de bandes floues, la longueur du trait vertical représente la demie-largeur de la bande. En trait plein la courbe théorique (éq. 3) relative à $h = 1{,}5$ Å; en trait interrompu celle relative à $h = 2{,}0$ Å.

○ phase cholestérique
● phase paracristalline
▲ phase complexe

(a) diméthylformamide
(b) pyridine
(c) *méta*-crésol (*v. infra*)

Noter la zone de démixtion dans (b) ($c \simeq 0{,}55$) et (c) ($c \simeq 0{,}40$).

lorsque la concentration s'élève. Nous avons porté dans la Figure 3 la position du centre de cette bande, en fonction de la concentration, et nous avons indiqué sa demi-largeur; pour quelques expériences nous avons représenté en outre (Fig. 4) la trace au microdensitomètre du profil de la bande.

Si on suppose que les molécules, rigides, sont organisées, du moins de manière approchée, selon un réseau hexagonal à deux dimensions (Robinson *et al.*, 1958), on peut admettre que la bande correspond à la première raie équatoriale de ce réseau: on peut alors exprimer (éq. 3) l'espacement de la bande en fonction de la concentration et de la masse linéaire des bâtonnets. Nous avons représenté, dans les Figures 3 et 4, s_1 en fonction de c, pour deux valeurs de la masse linéaire: celle de l'hélice α ($h = 1{,}5$ Å) et celle correspondant aux mesures en solution diluée ($h = 2$ Å, voir ci-dessus).

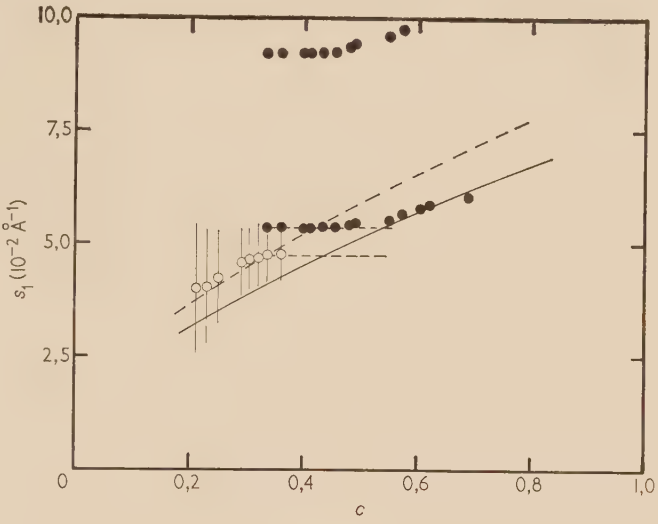


FIG. 3 (c)

Les résultats dépendent du solvant: *dans la pyridine la position de la bande est en bon accord avec la masse linéaire de l'hélice* α, *dans le* méta-*crésol, au contraire l'accord est bien meilleur avec la masse linéaire obtenue en solution diluée.*

Dans la diméthylformamide cette phase n'existe qu'à température élevée ($t > 40$°C environ): elle est absente à température ordinaire.

Il convient de remarquer que l'interprétation de la position d'une bande diffuse en termes de réseau cristallin ne peut pas être justifiée de manière tout à fait rigoureuse: dans les cas discutés ici, toutefois, l'évolution de la bande avec la concentration, et la différence de position avec le solvant, semblent justifier ce traitement simplifié.

(d) *Phase paracristalline*

Cette phase existe dans le système PLGB *méta*-crésol aux concentrations les plus élevées en polypeptide. Les preparations sont opaques, de consistance cireuse: au microscope polarisant la biréfringence est forte, sans qu'on observe de textures bien caractéristiques.

Dans la partie centrale des diagrammes de rayons X on observe quelques raies fines et intenses, dont les espacements se suivent dans l'ordre 1: $\sqrt{3}$: $\sqrt{4}$: $\sqrt{7}$: $\sqrt{9}$ typique d'un réseau hexagonal à deux dimensions: l'intensité des raies décroît très rapidement avec s croissant.

Nous avons porté dans la Figure 3(c) l'espacement de la première raie, s_1 en fonction de la concentration, ainsi que les courbes théoriques (3) relatives à un assemblage hexagonal de bâtonnets dont $h = 1{,}5$ Å et $h = 2{,}0$ Å. Les points expérimentaux sont en assez bon accord avec $h = 1{,}5$ Å, valeur correspondant à l'hélice α.

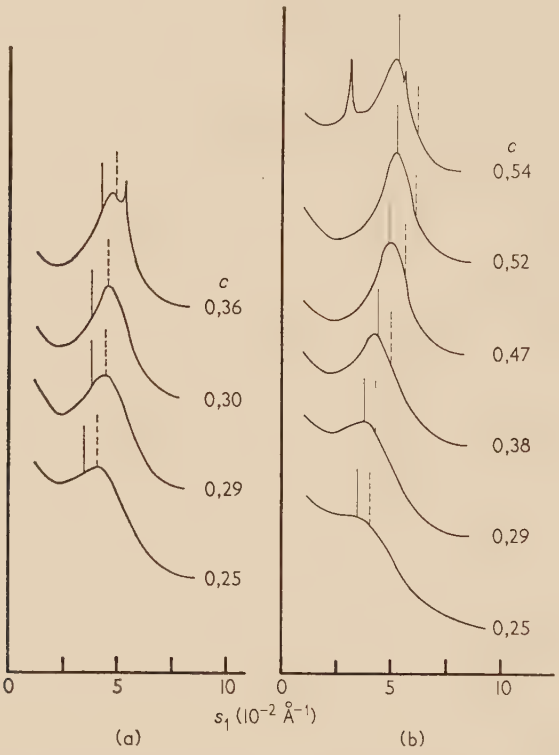


FIG. 4. Trace au microdensitomètre de la bande de la phase cholestérique. On a représenté par un trait plein vertical la position théorique (éq. 3) de la raie relative à $h = 1{,}5$ Å, par un trait interrompu celle relative à $h = 2{,}0$ Å.

(a) *méta*-crésol

$c = 0{,}25$: démixtion entre phases isotrope et cholestérique
$c = 0{,}36$: démixtion entre phases cholestérique et paracristalline

(b) pyridine

$c = 0{,}25$: démixtion entre phases isotrope et cholestérique
$c = 0{,}54$: démixtion entre phases cholestérique et complexe

Cette phase est donc formée d'un ensemble de bâtonnets rigides (vraisemblablement des hélices α) organisés selon un réseau hexagonal à deux dimensions, tout le solvant étant logé entre les molécules. Nous appellons cette phase "paracristalline", dénomination délibérément vague, car aux concentrations les plus élevées on observe, en plus des raies équatoriales du réseau hexagonal, plusieurs raies et bandes, indices d'une amorce d'organisation cristalline tridimensionnelle.

(e) *Phase complexe*

Nous avons rencontré cette phase avec deux solvants: la diméthylformamide et la pyridine. L'aspect au microscope polarisant est analogue à celui de la phase paracristalline.

Les diagrammes de diffraction des rayons X contiennent une famille de raies fines, dont les espacements sont ceux des raies équatoriales d'un réseau hexagonal. La phase complexe est caractérisée par les dimensions de la maille hexagonale, qui dépassent largement celles des mailles de la phase paracristalline (Fig. 3) et par l'intensité des raies qui n'est pas régulièrement décroissante, lorsque s croît, comme il en est dans la phase paracristalline (voir Appendice 2). En outre les dimensions du réseau ne varient pas avec la concentration: la phase complexe a donc une composition fixe. Cette phase est à bien des égards analogue à la phase que nous avons appelée "hexagonale complexe" des systèmes savon-eau (Luzzati *et al.*, 1960; Husson *et al.*, 1960).

L'exemple le plus frappant est fourni par la diméthylformamide. Aux faibles concentrations on n'observe, dans les diagrammes de rayons X, qu'une tache centrale diffuse; aux concentrations plus élevées (le seuil est situé aux environs de $c = 0{,}15$) le système de raies fines de la phase complexe apparaît et se superpose à la tache diffuse. L'intensité relative des raies, ainsi que leurs espacements, sont indépendants de la concentration: l'intensité de l'ensemble des raies par rapport à la tache diffuse augmente, au contraire, avec la concentration et la tache diffuse disparaît vers $c = 0{,}6$ (Fig. 3). Vers $c = 0{,}7$ les raies commencent à s'élargir, et un fond de diffusion continue apparaît, indice de désordre: aux concentrations les plus élevées la première raie s'estompe, et le diagramme de rayons X devient semblable à celui de la phase paracristalline des autres systèmes. Il est difficile, dans la diméthylformamide, de connaître la composition de la phase complexe, car on ne peut pas déterminer avec une précision suffisante la concentration à laquelle la tache de diffusion centrale a entièrement disparu.

Ce problème peut être résolu par l'étude des solutions dans la pyridine. Dans ce cas la phase complexe, qui présente toutes les caractéristiques décrites ci-dessus, est précédée, dans l'ordre des concentrations croissantes, par la phase cholestérique, avec laquelle on l'observe en démixtion (Fig. 3(b)): en étudiant avec soin la zone de transition on peut déterminer la concentration avec une assez bonne précision ($c = 0{,}55$).

En connaissant la concentration et les dimensions de la maille hexagonale (Appendice 2) on peut alors déterminer (éq. 3) la masse linéaire de l'ensemble du polypeptide contenu dans une maille élémentaire. On obtient $\mu = 225$ él. Å^{-1}, valeur plusieurs fois supérieure à celle des deux types de bâtonnets rencontrés dans les autres phases ($h = 1{,}5$ Å, $\mu = 77{,}5$ él. Å^{-1}; $h = 2{,}0$ Å, $\mu = 58$ él. Å^{-1}). La valeur élevée de μ, ainsi que l'intensité relativement faible de la première raie par rapport aux raies suivantes (voir Appendice 2), indiquent que la structure est plus complexe que celle des autres phases à symétrie hexagonale: puisque dans le cas de la pyridine la phase complexe est encadrée par deux structures formées d'hélices α (phase cholestérique et polypeptide sec), nous estimons qu'elle contient plusieurs hélices de ce type: le nombre est en fait trois, car la valeur de μ de la phase complexe est très voisine du triple de celle de l'hélice α.

Parmi les différentes manières selon lesquelles peuvent être associées ces trois hélices, la plus satisfaisante nous semble être un enroulement hélicoïdal autour d'un même axe, analogue à la structure d'une corde à trois torons (Fig. 5); c'est le type d'association

naturelle d'un petit nombre d'hélices parallèles de même sens, décrit par Crick (1953) sous la dénomination de "coiled-coil". Une vérification de ce modèle est fournie par l'accord satisfaisant qui existe entre les facteurs de structure observés, et ceux calculés pour la structure ayant les dimensions indiquées dans la Figure 5.†

Une autre vérification est de nature stéréochimique: la structure est compacte, les distances qui séparent les molécules voisines appartenant à une hélice secondaire ou à deux hélices contigües, sont approximativement les mêmes, sans qu'aucune distance ne devienne trop courte (Fig. 5 et Crick, 1953). Les différentes hélices secondaires

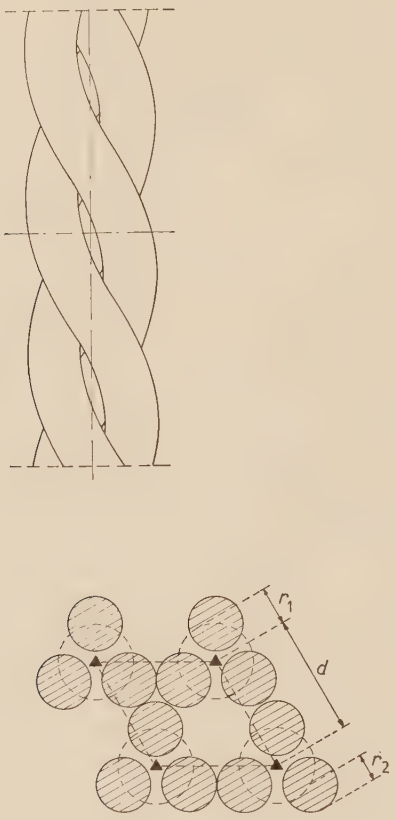


Fig. 5. Structure de la phase complexe.

s'encastrent vraisemblablement les unes dans les autres, de manière analogue à celle décrite par Franklin & Klug (1956), dans le cas du virus de la mosaïque du tabac, en formant une structure ordonnée dans les trois dimensions; cela peut expliquer la composition fixe en polypeptide et solvant, et justifie la présence, dans le clichés de rayons X, de plusieurs raies fines (et très peu intenses) dont les espacements ne correspondent pas aux noeuds équatoriaux du réseau hexagonal.

Il convient de remarquer toutefois que ce modèle, bien qu'il soit parfaitement satisfaisant, n'est pas le seul qui soit compatible avec les données expérimentales.

† Ayant tenu compte seulement des raies équatoriales, nous ne pouvons obtenir aucun renseignement sur le pas de l'hélice secondaire.

4. Conclusions et Discussion

Résumons brièvement les observations expérimentales. Les diagrammes des phases des trois systèmes sont récapitulés dans le Tableau 3. Dans tous les cas la transition

TABLEAU 3

Représentation schématique des diagrammes des phases

Entre parenthèses les valeurs de h (Å)

	isotrope	cholestérique	paracristalline	complexe
diméthylformamide	(1,96)	—	—	($\sim 3 \times 1{,}5$)
pyridine	(1,93)	($\sim$ 1,5)	—	($\sim 3 \times 1{,}5$)
méta-crésol	(1,88)	($\sim$ 2)	($\sim$ 1,5)	—

entre phases se fait par une zone de démixtion (Fig. 3), en accord avec la règle des phases. Aux concentrations les plus élevées ($c > 0{,}7$) les structures deviennent désordonnées: on se trouve probablement en présence de solutions solides des phases paracristalline ou complexe avec le polypeptide pur.

Dans toutes les phases le polypeptide a le comportement d'un bâtonnet rigide, dont on peut déterminer la masse linéaire: nous avons caractérisé ce paramètre par la distance h entre deux groupes peptidiques contigus, projetée sur l'axe du bâtonnet. Les valeurs de h sont portées dans le Tableau 3: elles se groupent autour de 1,5 Å et de 2 Å. Dans les solutions isotropes $h \sim 1{,}93$ Å: la transition vers la forme ayant $h = 1{,}5$ Å coïncide avec le passage de la solution isotrope à une phase organisée, dans les cas de la diméthylformamide et de la pyridine, et avec le passage de la phase cholestérique à la phase paracristalline, dans le cas du *méta*-crésol.

La connaissance de h, et éventuellement de R_c, ne suffit pas, pour déterminer la structure des bâtonnets. Mais les présomptions en faveur des structures hélicoïdales sont très fortes: or dans les polypeptides les configurations hélicoïdales stables n'existent qu'en nombre fini, et petit. Donohue (1953) a essayé d'estimer la stabilité relative de ces structures, en analysant les contraintes dues à la déformation des liaisons chimiques, à l'intérieur du squelette peptidique:

Type de structure	h	Instabilité par résidu, par rapport à l'hélice α
ruban α_{II}	2,75 Å	0,5 kcal
hélice 3_{10}	2,00	1,0
hélice α	1,50	0,0
hélice π	1,15	0,5
hélice 4, 3_{13}	1,20	2,4
hélice γ	0·98	2,0

Il est remarquable que les deux valeurs de h que nous avons rencontrées, correspondent à deux de ces configurations, parmi les plus stables: l'hélice α et l'hélice 3_{10}. Il est peu probable que cette coïncidence soit fortuite: le PLGB adopte vraisemblablement ces deux structures, dans les conditions où nous l'avons étudié.

Puisque l'énergie des squelettes peptidiques est presque la même dans ces deux structures, ce sont les interactions avec le solvant qui déplacent l'équilibre dans le sens d'une des deux configurations. On peut supposer, par exemple, que la flexibilité des chaînes latérales dépend de la nature du solvant. Lorsque la concentration s'élève, les chaînes latérales des différentes molécules viennent en contact d'autant plus facilement qu'elles sont plus étirées; dès que ce contact s'établit, le noyau peptidique est perturbé et il se crée des conditions favorables à un changement de structure. C'est également dans les interactions entre le polypeptide et le solvant qu'il faut chercher l'explication des différences entre les diagrammes des phases des différents systèmes. Pour préciser ce mécanisme il serait utile de connaître le rôle de la polarité du solvant sur la structure des hélices: nous espérons que l'étude des systèmes PLGB dioxanne et PLGB benzène, actuellement en cours, nous aidera à résoudre ce problème.

La phase complexe offre un exemple de structure torsadée (coiled-coil) qui confirme les idées avancées par Crick (1953) sur des bases théoriques. Par ailleurs il est intéressant de constater que la phase cholestérique peut être formée d'hélices de type différent, selon la nature du solvant.

Il est peut-être prématuré de chercher des implications biologiques de ces observations. Il est toutefois remarquable que le seul changement de la concentration puisse perturber la structure d'un polypeptide en modifiant l'hélice fondamentale, et en faisant apparaître une variété de structures organisées. On peut ainsi se demander si d'une part la description de la structure des protéines en termes d'hélice α n'est pas un peu trop schématique (Doty, 1959), et si d'autre part des transitions entre formes hélicoïdales ne pourraient accompagner certaines modifications morphologiques.

Pour terminer nous essayerons d'expliquer la contradiction flagrante qui résulte de la confrontation de nos résultats avec ceux d'autres travaux, d'après lesquels, en effet, le PLGB devrait posséder dans les conditions de nos expériences, la structure de l'hélice α†: il convient donc d'examiner ces travaux, et d'en analyser les interprétations. Les expériences peuvent être groupées en trois catégories.

(a) *Diffusion de la lumière*

Dans le cas du PLGB cette technique permet de déterminer la masse moléculaire M et le rayon de giration R_0‡. Si les molécules ont la forme de bâtonnets rigides, R_0 est proportionnel à M; on peut alors déterminer h

$$h = \frac{219 \times R_0 \times \sqrt{12}}{M} \tag{4}$$

qui est indépendant de M. Plusieurs auteurs ont étudié le PLGB par diffusion de la lumière (Doty *et al.*, 1956; Wada, 1960; Spach, 1960), et ont vu dans les résultats expérimentaux une preuve de l'existence d'hélices α. Toutefois en analysant les résultats de ces expériences (Fig. 6) on remarque que s'il est vrai que la valeur moyenne du rapport (4) est voisine de 1,5 Å, en fait elle varie systématiquement avec M, en diminuant pour M croissant. Cette variation, constatée d'ailleurs par ces auteurs, est un indice de flexibilité des bâtonnets; pour déterminer correctement dans ce cas la

† Les concentrations sont en fait un peu plus élevées dans les expériences de rayons X que dans les autres études. Nous avons vérifié, par diffusion de la lumière, qu'aucune anomalie n'existe aux concentrations intermédiaires.

‡ Lorsqu'il s'agit de bâtonnets polydispersés, dont $f(M)$ est la distribution des masses, en poids, on mesure en fait des valeurs moyennes: $M_w = \int Mf(M)\mathrm{d}M$ et $< R_0 > = [\int MR^2{}_0(M)\mathrm{f}(M)\mathrm{d}M / \int Mf(M)\mathrm{d}M]^{\frac{1}{2}}$

valeur de h il faut extrapoler le rapport (4) vers $M = 0$. Cette extrapolation est assez incertaine, à cause de la dispersion des points expérimentaux: il est clair, toutefois (Fig. 6) qu'elle conduit à une valeur de h voisine de 2 Å, ou même supérieure. Il convient de remarquer que l'effet de la polydispersité tend à augmenter la valeur de (4): toutefois quelques mesures récentes (Spach, 1960, et travaux en cours au Centre de Recherches sur les Macromolécules) indiquent que la correction est comprise entre 5 et 15%, donc insuffisante pour rendre compte de l'écart entre les deux valeurs de h 1,5 à 2 Å).

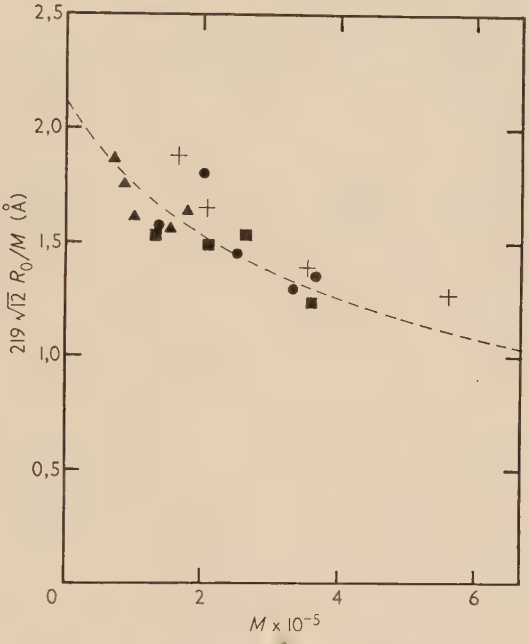


FIG. 6. Représentation des résultats de diffusion de la lumière.

+ PLGB brut dans la diméthylformamide (Spach, 1960).
● PLGB fractionné dans la diméthylformamide (Spach, 1960).
■ PLGB dans un mélange chloroforme-formamide (Doty *et al.*, 1956).
▲ PLGB dans le dichloroéthane (Wada, 1960).

On pourrait essayer d'interpréter la distribution des points experimentaux (Fig. 6) par un modèle plus élaboré que le bâtonnet rigide, par exemple la "worm-like chain" de Porod (1949), qui peut convenir pour décrire la configuration d'un bâtonnet lorsque celui-ci devient légèrement flexible. Dans ce modèle la rigidité de la molécule est caractérisée par un paramètre q, dit "longueur de persistance". L'expression de (4) est dans ce cas (Benoit & Doty, 1953):

$$\frac{R_0 \times 219 \times \sqrt{12}}{M} = h\sqrt{12}\left[\frac{1}{3x} - \frac{1}{x^2} + \frac{2}{x^3} - \frac{2}{x^4}(1 - e^{-x})\right]^{\frac{1}{2}} \qquad (5)$$

où $x = \dfrac{Mh}{219q}$. Nous avons tracé dans la Figure 6 la fonction (5) pour $q = 505$ Å et $h = 2{,}12$ Å, la valeur de h étant 10% plus élevée que celle obtenue par rayons X, pour tenir compte, du moins en première approximation, des effets de la polydispersité. L'accord des points expérimentaux avec la courbe théorique est assez bon, bien meilleur, en tout cas, qu'avec la droite horizontale $h = 1{,}65$ Å, qui correspondrait à l'hélice α (avec la même correction de polydispersité).

(b) *Techniques hydrodynamiques*

Qu'il s'agisse de mesures de viscosité (Doty *et al.*, 1956; Yang, 1958*a*; Spach, 1960), de biréfringence d'écoulement (Yang, 1958*b*), de diffusion de translation (Spach, 1960; Freund, 1961), de diffusion de rotation (Tinoco, 1957, O'Konski *et al.*, 1959; Marchal & Marchal, 1959; Benoit & Boeckel, en préparation), l'interprétation des expériences en termes de dimensions moléculaires fait appel à un certain nombre d'hypothèses:

(a) la particule est parfaitement rigide,

(b) la structure est compacte et impénétrable au solvant,

(c) les lois de l'hydrodynamique des milieux continus sont encore valables à l'échelle des dimensions moléculaires,

(d) la forme de la particule peut être décrite par un ellipsoïde ou un cylindre de révolution.

Sans entrer dans l'analyse détaillée de ces hypothèses (la première au moins, ne semble pas être tout à fait en accord avec les expériences de diffusion de la lumière) il est évident que leur accord avec le comportement réel des molécules de PLGB ne peut être qu'approximatif: il nous paraît très douteux que l'interprétation de ces expériences puisse permettre de choisir entre $h = 2$ Å et $h = 1{,}5$ Å. Ce problème est d'ailleurs à l'étude au Centre de Recherches sur les Macromolécules (Freund, 1961).

(c) *Techniques optiques*

Nous incluons dans cette denomination des techniques comme l'absorption dans l'infra-rouge (Bamford *et al.*, 1956), le pouvoir optique rotatiore (Yang & Doty, 1957; Downie *et al.*, 1957), les moments diélectriques (Wada, 1960), qui, dans le cas des polypeptides, permettent de reconnaître les configurations de type α. Telles qu'on les a employées jusqu'ici, ces techniques ne semblent pas être en mesure de discerner entre deux formes hélicoïdales (hélices α et 3_{10}) dont la structure est très voisine.

Si ce travail a pu être mené à bien, c'est grâce à l'intérêt que lui ont porté plusieurs de nos collègues. Nous tenons à remercier tout particulièrement MM. H. Benoit, M. Daune, L. Freund, et Mme. E. Marchal.

L'un de nous désire exprimer sa reconnaissance à Olin Mathieson Chemical Corp. qui, en lui accordant une bourse, lui a permis de participer à ce travail, et remercie plus particulièrement le Dr. A. Pace, Directeur de Recherches de cette Compagnie qui a toujours suivi cette étude avec le plus vif intérêt.

APPENDICE 1

Variation de $j_n(s)$ avec la Masse Moléculaire du Polypeptide

On constate (Tableau 2) que A/c_e est légèrement plus grand pour les échantillons de masse élevée que pour ceux de masse faible; le modèle d'un ensemble de bâtonnets longs et rigides, indépendants entre eux, ne peut pas rendre compte de ce phénomène.

On peut parfaire le modèle en admettant l'existence de certains accidents: nous nous sommes occupés de ce problème (Luzzati & Benoit, 1961), en calculant les effets des extrémités libres des bâtonnets, et des points d'articulation et de croisement entre bâtonnets. Nous avons montré, dans ce cas, que l'intensité diffusée reste égale à celle correspondant aux bâtonnets longs et indépendants, lorsque s est suffisamment grand; au voisinage de l'origine, au contraire, il convient d'ajouter un terme correctif

$$j_n(s) = A\tfrac{1}{2} \exp(-\pi^2 R_c^2 s^2) K_0(\pi^2 R_c^2 s^2) + Bu(s) \qquad (6)$$

La forme de ce terme $u(s)$ est connue, et le coefficient B est la somme des contributions, de signe négatif, des extrémités libres, et de celles, positives, des points d'articulation et de croisement.

Dans tous les échantillons étudiés ici la masse moléculaire est suffisamment élevée pour que les extrémités libres aient un effet négligeable: B ne peut donc pas être négatif, ce qui enlève à A/c_e toute raison de pécher par défaut dans les échantillons de faible masse. Il est vraisemblable, au contraire, que dans les échantillons de masse élevée les bâtonnets ne soient pas parfaitement rigides (voir discussion relative à la diffusion de la lumière) où même que les manipulations qu'ont subies les solutions créent des enchevêtrements, inextricables sous le seul effet de l'agitation thermique: il faut indiquer, en effet, que les solutions, qui sont très visqueuses, surtout lorsque la masse du polypeptide est élevée, sont injectées à l'intérieur des cuves porte-échantillon sous une forte pression, à travers un tube capillaire dont le diamètre est de 0,5 mm et la longueur de 5 mm. Dans un cas comme dans l'autre B est positif, et, en interprétant les courbes expérimentales $j_n(s)$ comme si les bâtonnets étaient rigides et indépendants on obtient pour A/c_e des valeurs apparentes trop élevées.

APPENDICE 2

Quelques Vérifications Cristallographiques de la Structure de la Phase Complexe

Calculons le facteur de structure $F(s)$ dans le plan équatorial, d'un élément formé par l'association de trois bâtonnets qui s'enroulent en hélice autour d'un même axe (Fig. 5). Si r_1 est la distance du centre des bâtonnets à l'axe de l'helice, et $f(s)$ est le facteur de structure d'un des bâtonnet isolé et étiré, on a:

$$F(s) = f(s)J_0(2\pi r_1 s) \tag{7}$$

Si les bâtonnets ont une forme de cylindres compacts de rayon r_2

$$f(s) = \frac{J_1(2\pi r_2 s)}{s} \tag{8}$$

$$F(s) = 1/s \,.\, J_0(2\pi r_1 s) J_1(2\pi r_2 s) \tag{9}$$

On peut déterminer le rayon des bâtonnets, si on admet que $\mu = 77{,}5$ él. Å^{-1}, comme dans l'hélice α, et $\psi = 2{,}469$ Å^3. él^{-1}

$$\pi r_1^2 = \mu\psi = 190 \quad r_2 = 7{,}8\text{Å}.$$

Nous avons mesuré l'intensité des raies par simple estimation visuelle; nous avons calculé F_{obs} en divisant l'intensité par un facteur de multiplicité et par un facteur géométrique s^{-2}, valable aux petits angles.

Nous avons effectué le calcul de $F(s)$ avec $r_2 = 7{,}8$ Å, et pour différentes valeurs de r_1: c'est avec $r_1 = 10{,}9$ Å que nous avons obtenu le résultat le plus satisfaisant. L'accord entre F_{obs} et F_{calc} pourrait être amélioré, en faisant une hypothèse moins grossière sur la structure des bâtonnets.

h	k	Diméthylformamide s 10^{-2}Å^{-1}	$\lvert F_o \rvert$	$\lvert F_c \rvert$	Pyridine s 10^{-2}Å^{-1}	$\lvert F_o \rvert$	$\lvert F_c \rvert$
1	0	3,29	10	10	3,19	10	10
1	2	5,70	21	20	5,52	12	15
2	0	6,58	11	9	6,38	9	7,5
2	1	8,70	5	2			5

REFERENCES

Ambrose, E. J. & Elliott, A. (1951). *Proc. Roy. Soc.* A, **205**, 47.

Bamford, C. H., Elliott, A. & Hanby, W. E. (1956). In *Synthetic Polypeptides*. New York: Academic Press.

Benoit, H. & Doty, P. (1953). *J. Phys. Chem.* **57**, 958.

Blout, E. R. & Karlson, R. H. (1956). *J. Amer. Chem. Soc.* **78**, 941.

Crick, F. H. C. (1953). *Acta Cryst.* **6**, 689.

Donohue, J. (1953). *Proc. Nat. Acad. Sci., Wash.* **39**, 470.

Doty, P. (1959). *Rev. Mod. Phys.* **31**, 107.

Doty, P., Bradbury, J. H. & Holtzer, A. M. (1956). *J. Amer. Chem. Soc.* **78**, 947.

Downie, A. R., Elliott, A., Hanby, W. E. & Malcolm, B. R. (1957). *Proc. Roy. Soc.* A, **242**, 325.

Franklin, R. E. & Klug, A. (1956). *Biochim. biophys. Acta*, **19**, 403.

Freund, L. (1961). Thesis at the University of Strasbourg.

Husson, F., Mustacchi, H. & Luzzati, V. (1960). *Acta Cryst.* **13**, 668.

Low, B. W. (1953). In *The Proteins*, Vol. 1. New York: Academic Press.

Luzzati, V. (1960). *Acta Cryst.* **13**, 939.

Luzzati, V. & Benoit, H. (1961). *Acta Cryst.* **14**, 297.

Luzzati, V., Mustacchi, H., Skoulios, A. & Husson, F. (1960). *Acta Cryst.* **13**, 660.

Luzzati, V. & Nicolaïeff, A. (1959). *J. Mol. Biol.* **1**, 127.

Luzzati, V., Nicolaïeff, A. & Masson, F. (1961). *J. Mol. Biol.* **3**, 185.

Luzzati, V., Witz, J. & Nicolaïeff, A. (1961*a*). *J. Mol. Biol.* **3**, 367.

Luzzati, V., Witz, J. & Nicolaïeff, A. (1961*b*). *J. Mol. Biol.* **3**, 379.

Marchal, E. & Marchal, J. (1959). *Arch. Sci., Geneve, Fasc. Spéc.* **12**, 24.

Mitchell, J. C., Woodward, A. E. & Doty, P. (1957). *J. Amer. Chem. Soc.* **79**, 3955.

O'Konski, C. T., Yoshioda, K. & Orttung, W. H. (1959). *J. Phys. Chem.* **63**, 1558.

Pauling, L., Corey, R. B. & Branson, H. R. (1951). *Proc. Nat. Acad. Sci., Wash.* **37**, 205.

Perutz, M. F. (1951). *Nature*, **167**, 1053.

Porod, G. (1949). *Monatsh.* **80**, 251.

Robinson, C. (1956). *Trans. Faraday Soc.* **52**, 571.

Robinson, C., Ward, J. C. & Beevers, R. B. (1958). *Disc. Faraday Soc.* **25**, 29.

Spach, G. (1960). Thesis at the University of Strasbourg.

Tinoco, I. (1957). *J. Amer. Chem. Soc.* **79**, 761.

Yang, J. T. (1958*a*). *J. Amer. Chem. Soc.* **80**, 1783.

Yang, J. T. (1958*b*). *J. Amer. Chem. Soc.* **80**, 5139.

Yang, J. T. & Doty, P. (1957). *J. Amer. Chem. Soc.* **79**, 761.

Wada, A. (1960). *Bull. Chem. Soc. Japan*, **33**, 822.

Watson, G. N. (1952). In *A Treatise on the Theory of Bessel Functions*, Cambridge University Press.

J. Mol. Biol. (1961) **3,** 585–594

Thermal Renaturation of Deoxyribonucleic Acids †

Julius Marmur‡ and Paul Doty

Department of Chemistry, Harvard University, Cambridge, Massachusetts, U.S.A.

(*Received 29 March 1961*)

The conditions that produce the optimal re-formation of the native DNA conformation from denatured DNA have been examined. The restoration of the native conformation, called thermal renaturation, has been found to depend markedly on the source of the DNA; mammalian DNA coming from cells with very large DNA content renatures only slightly, bacterial DNA with greatly reduced DNA content per cell undergoes extensive renaturation, and the very smallest bacteria together with bacteriophage, having the lowest DNA contents, show nearly complete renaturation. With a given DNA, the optimal renaturation was found to occur at about 25° below the denaturation temperature, T_m. The extent of renaturation was optimal above 0·4 M-Na^+ and increased with molecular weight. The identity of the renatured DNA and the native material can be shown in two ways: the similarity of the absorbance-temperature curves and the similarity of the rate of thermal inactivation of biological markers at temperatures somewhat above T_m. This reproduceability of the helix-coil transition and the course of thermal inactivation demonstrates that the same secondary structure has re-formed and that non-specific hydrogen bonding is not involved.

1. Introduction

When DNA in aqueous solution is exposed to temperatures above the melting temperature, T_m (Marmur & Doty, 1959), the two strands in each molecule separate and assume randomly coiled configurations. Upon rapid cooling this denatured state is maintained, although considerable non-specific base pairing through hydrogen bonds occurs. However, if the thermally denatured DNA solution is cooled slowly, renaturation of the DNA takes place, resulting in a restoration of its transforming activity, physical, chemical and immunological properties (Marmur & Lane, 1960; Doty, Marmur, Eigner & Schildkraut, 1960; Levine, Murakami, Van Vunakis & Grossman, 1960).

The present communication is an attempt to describe the optimum conditions for the renaturation of DNA and to compare some properties of renatured and native DNA not previously described.

2. Materials and Methods

Transformation techniques using *Diplococcus pneumoniae* as well as the method of obtaining the absorbance-temperature profiles have been previously described (Fox & Hotchkiss, 1957; Marmur & Lane, 1960; Doty *et al.*, 1960; Doty, Boedtker, Fresco,

† This work was supported by a grant from the United States Public Health Service (C-2170).

‡ Present address: Graduate Department of Biochemistry, Brandeis University, Waltham, Massachusetts, U.S.A.

Haselkorn & Litt, 1959). DNA was isolated from the following organisms according to the method of Marmur (1961): *D. pneumoniae* (R-36A); *Bacillus subtilis* (168); *Escherichia coli* (K-12 and B); *Pseudomonas aeruginosa* (NRRL B-23); the pleuropneumoniae-like organism, *Mycoplasma gallisepticum* (avian PPLO 5969); and from tobacco leaves. DNA samples isolated from salmon sperm and calf thymus were a gift from Dr. N. Simmons. DNA from the T-even *E. coli* bacteriophages was generously supplied by Dr. H. Van Vunakis. The DNA base compositions were obtained from published reports (Belozersky & Spirin, 1960; Chargaff, 1955) and confirmed by the T_m of native samples (Marmur & Doty, 1959, 1961).

The molecular weight of the native DNA samples used in these studies was approximately $10 \pm 2 \times 10^6$. Sonic degradation was carried out in the presence of the free-radical trapping agent AET (aminoethyl *iso*thiuronium bromide hydrogen bromide) as described previously (Litt, Marmur, Ephrussi-Taylor & Doty, 1958).

3. Experimental

(a) *Effect of DNA source on renaturation*

Although the T_m for strand separation of DNAs having similar base compositions is independent of its source, the extent to which renaturation proceeds is greatly influenced by the type of DNA being studied. This becomes understandable when it is recognized that renaturation will depend on the concentration of *specific* complementary strands during the renaturation (Marmur & Lane, 1960). If there is no duplication of sequences in the DNA molecules in a given cell and if the molecular weight of the isolated DNA is approximately the same, it is expected that the concentration of complementary pairs will fall off with increasing complexity of the species from which the DNA is isolated. Bacteria, which possess a relatively small number of nucleic acid molecules (Vendreley, 1958), contain a more homogeneous distribution of DNA molecules than that of DNA isolated from higher plants or animals but less homogeneous than the DNA isolated from bacteriophages (Crampton, Lipshitz & Chargaff, 1954; Brown & Brown, 1958). Thus, for a given concentration of DNA, renaturation would be expected to occur most readily with bacteriophage DNA and to be least complete with DNA from mammalian source such as calf thymus or salmon sperm.

In order to explore this matter, renaturation was studied by following the absorbance at 260 mμ in the Beckman spectrophotometer. When DNA is denatured and rapidly cooled (quenched), approximately 60% of the bases are hydrogen bonded in a non-specific fashion, easily disrupted by exposure to temperature below the T_m (Doty *et al.*, 1959*a*). Upon raising the temperature to the region below the T_m, annealing conditions are provided. These weak intramolecular hydrogen bonds are observed to rupture, by the *increase* in absorbance, and this as well as the elevated temperature allows the complementary strands the time and the opportunity to renature into double-stranded molecules which are hydrogen bonded and in register. If renaturation occurs, the absorbance at 260 mμ will fall. Experiments of this type are shown in Fig. 1. It can be seen that after the melting of the weak, intramolecular hydrogen bonds of the denatured, fast-cooled DNA, renaturation of some of the DNA sample occurs. The rate and extent of renaturation of DNA is seen to be dependent upon its source. Calf thymus DNA (as well as DNA from salmon sperm and tobacco leaves) does not renature; *D. pneumoniae* DNA (as well as DNA from most bacterial sources) readily renatures, whereas the DNAs of T6r^+ (as well as T2 and T4) and *Mycoplasma gallisepticum* (avian PPLO), which are more homogeneous (Hershey & Burgi, 1960; Guild, Morowitz & Castro, 1960; Guild, 1961), renature the most readily.

(b) *Optimum renaturation temperature: biological studies*

Renaturation of DNA by slowly cooling a sample in a large water bath has led to somewhat variable results owing to the difficulty of controlling accurately the rate of cooling. In order to find the optimum renaturation temperature of *D. pneumoniae* DNA, denatured samples were exposed at different constant temperatures for varying periods of time. The rate and extent of renaturation was followed either by the recovery of transforming activity or by the decrease in absorbance.

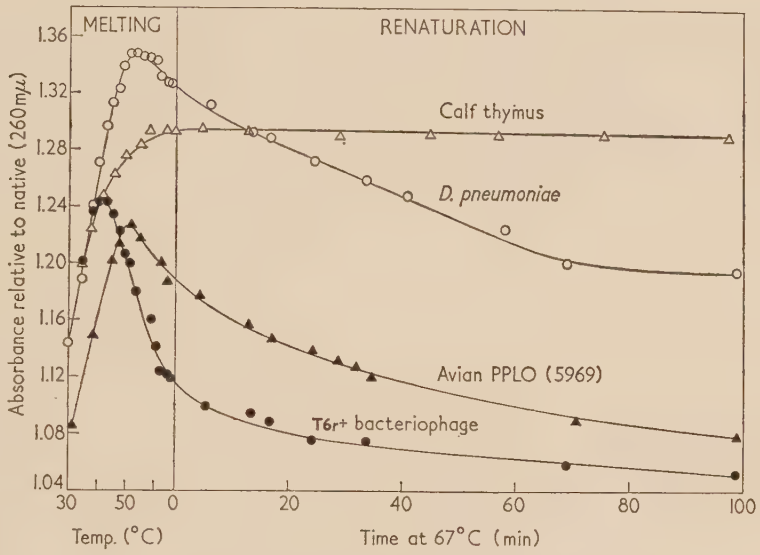


Fig. 1. Thermal transitions of denatured DNA samples.

DNA at 20 μg/ml. was heated at 100°C for 10 min in 0·3 M-NaCl plus 0·03 M-Na citrate and quickly cooled. The samples of DNA (in ground glass stoppered cuvettes) at the same concentration and in the same solvent, were then placed in the Beckman spectrophotometer chamber, prewarmed to 67°C. The initial part of the graph represents the *increase* in absorbance at 260 mμ as the denatured DNA samples reach temperature equilibrium. The curves to the right of the vertical line represent the *decrease* in absorbance of the temperature-equilibrated, renaturing samples as a function of time of exposure at 67°C. DNA was isolated from the organisms listed in the Figure.

When the course of renaturation of *D. pneumoniae* DNA is followed by the restoration of the ability to transform with respect to the streptomycin resistance marker, the results are found to depend very much on the temperature. As shown in Fig. 2, the recovery of biological activity displays an initial rapid rate, followed by a slower one. At the optimum temperature, the slow increase continues for at least 5 hr. When a plateau is reached in the restoration of the biological activity, subsequent cooling to room temperature results in an additional increment in transforming activity (10 to 50%) depending on the previous temperature. By exposing denatured DNA to the optimum renaturation temperature for 2 to 3 hr, as well as a subsequent slow cooling to room temperature, restoration of 40 to 50% of the original biological activity is generally obtained in the case of *D. pneumoniae* DNA. These conditions may readily be used for the renaturation of DNA which has suffered denaturation by means other than thermal exposure, such as formamide (Marmur & Ts'o, 1961); thus a mild method for renaturation, which avoids exposure to relatively high temperature, is available.

In order to find the optimum temperature for the restoration of the transforming activity of *D. pneumoniae* DNA at the ionic strength employed, the initial rates of

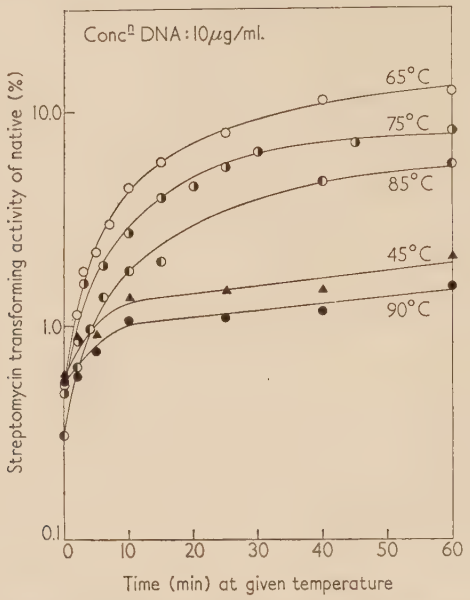


FIG. 2. Effect of temperature on the recovery of streptomycin-transforming activity of thermally denatured *D. pneumoniae* DNA.

DNA from streptomycin resistant cells of *D. pneumoniae* was denatured at 10 μg/ml. by heating for 10 min at 100°C in 0·3 M-NaCl plus 0·03 M-Na citrate. Portions, at the same DNA and solvent concentration, were then distributed and exposed to various temperatures. Samples were removed and quickly cooled at different times and assayed for streptomycin-transforming activity.

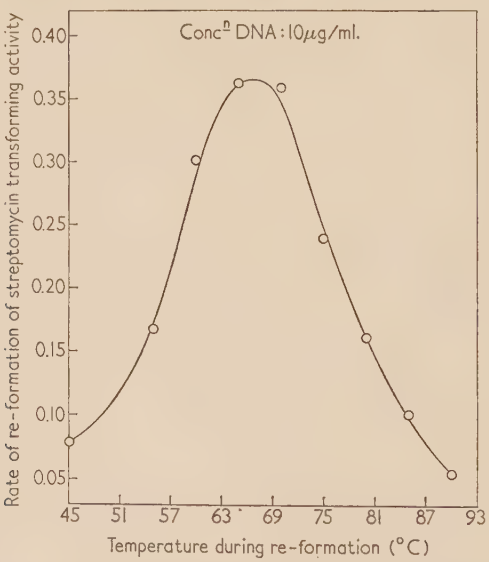


FIG. 3. Effect of temperature on the recovery of transforming activity of denatured *D. pneumoniae* DNA in 0·3 M-NaCl plus 0·03 M-Na citrate: biological study. The initial rates of recovery of the biological activity of the DNA shown in Fig. 2 are plotted against the renaturation temperature.

recovery of the biological activity were plotted (Fig. 3) against the renaturation temperature. The optimum temperature lies within the range 65 to 70°C. The slightly higher value obtained by this method, compared with that obtained by the absorbance method (see following section), may be a reflection of the observation that the streptomycin resistance marker (used in the biological assay for the renaturation) is probably associated with a DNA molecule which has a higher guanine *plus* cytosine content than the average composition of the total DNA population (Marmur & Lane, 1960).

(c) *Optimum renaturation studies: absorbancy studies*

It has been shown (Fig. 1) that the renaturation of DNA can be followed by the decrease in absorbance after the intramolecular hydrogen bonds of denatured, fast-cooled DNA are melted out. When the initial rate of decrease in the absorbance is plotted against the renaturation temperature, the curve shown in Fig. 4 is obtained. The optimum renaturation temperature for *D. pneumoniae* DNA is 65°C.

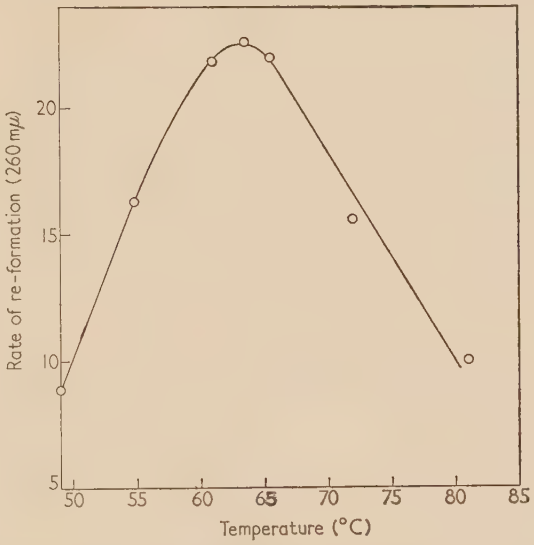


FIG. 4. Effect of temperature on the renaturation rate of denatured *D. pneumoniae* DNA in 0·3 M-NaCl plus 0·03 M-Na citrate: absorbancy study. Conditions are similar to those described under Fig. 1. The initial rates of *decrease* in absorbance after temperature equilibrium is reached are plotted against the renaturation temperature.

(d) *Effect of base composition on the optimum renaturation temperature*

Using the absorbance method, the optimum renaturation temperature for DNA from various bacterial sources was determined. When the optimum temperature was plotted against the guanine *plus* cytosine content of the DNA, the results shown in Fig. 5 were obtained. As the guanine *plus* cytosine content increases, there is a gradual increase in the optimum renaturation temperature.

(e) *Effect of molecular weight*

In order to compare the rate and extent of the renaturation of denatured DNA as a function of the molecular weight, sonically degraded *D. pneumoniae* DNA of molecular

weight $0{\cdot}55\times10^6$ was compared with DNA of molecular weight 10×10^6. The degraded material retained 0·24% of its transforming activity with respect to the streptomycin resistance marker. The renaturation of thermally denatured samples of both the

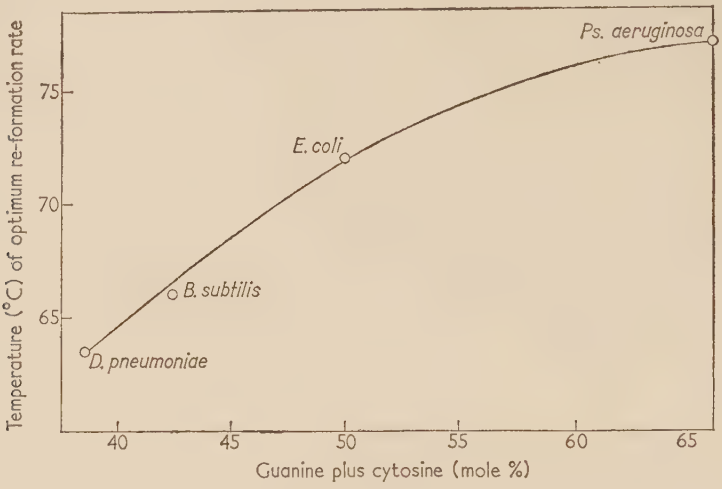


Fig. 5. Effect of guanine *plus* cytosine content on the optimum renaturation temperature of denatured DNA in 0·3 M-NaCl plus 0·03 M-Na citrate. DNA from *D. pneumoniae*, *Bacillus subtilis*, *E. coli* and *Pseudomonas aeruginosa* were denatured, quickly cooled and adjusted to 0·3 M-NaCl plus 0·03 M-Na citrate. The optimum renaturation temperature was then determined as described below Fig. 4 and then plotted against the guanine *plus* cytosine content of the respective bacterial DNA.

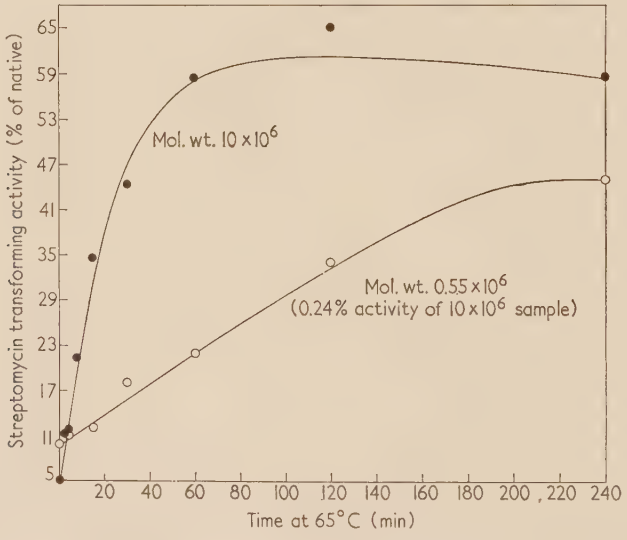


Fig. 6. Effect of molecular weight on the recovery of streptomycin transforming activity of denatured *D. pneumoniae* DNA in 0·3 M-NaCl plus 0·03 M-Na citrate. Conditions are described in the text and are similar to those employed for Fig. 2.

undegraded and sonically treated DNA samples was studied by the restoration of the biological activity under optimum conditions of temperature and ionic strength.

In Fig. 6 are shown the differences in the rate and extent of renaturation of the two samples. The transforming DNA was first denatured by heating the samples in

0·3 M-NaCl plus 0·03 M-Na citrate at 100°C for 10 min and then incubated at 65°C at 15 μg per ml. in the same solvent. Samples, removed at various times and quickly cooled, were assayed for the ability to transform sensitive cells of *D. pneumoniae* to streptomycin resistance. It can readily be seen that both the rate and extent of the recovery of biological activity are dependent on the molecular weight. Similar results are obtained when renaturation is followed by the absorbance-temperature method.

Experiments summarized in Fig. 7 again clearly show that the re-formation is more rapid and extensive the higher the molecular weight. In this experiment, heat denatured, fast-cooled samples in ground glass stoppered cuvettes were placed in the

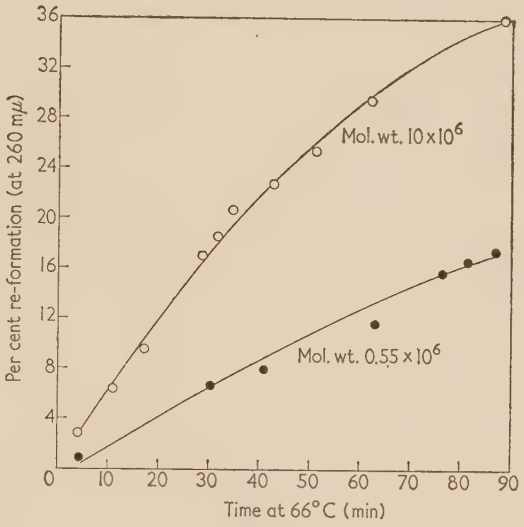


Fig. 7. Effect of molecular weight on the renaturation of denatured *D. pneumoniae* DNA in 0·3 M-NaCl plus 0·03 M-Na citrate. The samples were the same as used in Fig. 6. Conditions are described in the text and are similar to those employed for Fig. 1.

Beckman spectrophotometer chamber prewarmed to 66°C. When the intramolecular hydrogen bonds had melted, as measured by the *increase* in absorbance, the renaturation, indicated by the subsequent *decrease* in absorbance, was followed for both samples. After maximum re-formation had occurred, the percentage re-formation was plotted as a function of time of exposure to 66°C.

(f) *Thermal stability of renatured DNA*

If renaturation is indeed the restoration of hydrogen bonding between complementary strands whose bases are in register, then the thermal stability of renatured DNA should mimic that of native DNA. This was tested by exposing both renatured and native DNA to a temperature several degrees above the T_m and following their transforming activity as a function of time. This has been shown previously (Doty, Marmur & Sueoka, 1959) to result in an easily measurable, time-dependent loss of transforming activity due to denaturation of the molecules carrying the genetic marker.

Renatured DNA was prepared as follows: *D. pneumoniae* was heated at 100°C for 10 min in 0·3 M-NaCl plus 0·03 M-Na citrate and then exposed to 65°C at 20 μg/ml. in the same solvent for varying periods of time. Samples were removed at several times, quickly cooled, and dialysed against 0·15 M-NaCl plus 0·015 M-Na citrate.

The thermal inactivation curves at 89·2°C in 0·15 M-NaCl *plus* 0·015 M-Na citrate of the native and renatured dialysed samples are shown in Fig. 8. DNA renatured for 3 hr (41% restoration of biological activity) as well as samples renatured for 25 min (33% restoration) and 60 min (36% restoration) showed the same thermal stability as did native DNA. Data for the latter two renatured samples are not shown in the Figure. DNA renatured for 4 min (7% restoration of biological activity) shows an abnormally higher thermal resistance which may be a reflection of the presence of a resistant component selected during the initial (100°C) thermal exposure.

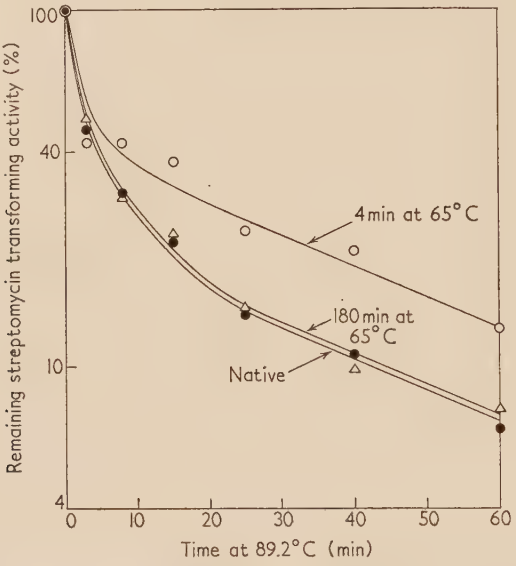


FIG. 8. Stability of renatured and native *D. pneumoniae* DNA to thermal inactivation of streptomycin transforming activity. Conditions are described in text.

It is suggested that this (or the absorbance-temperature profile) criterion of the thermal stability of the material should be used as a test of its native-like structure to rule out the possibility of non-specific aggregation.

The absorbance-temperature profile can also be employed as a criterion to check the structure of renatured DNA. The profiles of native and denatured DNA differ significantly (Doty *et al.*, 1959*a*). Whereas native DNA melts very sharply, resulting in a steep rise in optical density in the T_m region, denatured DNA gives rise to a broad profile.

Because of the high degree of homogeneity of the DNA from phage and PPLO, an attempt was made to see to what extent DNAs from these sources could be renatured and to examine their thermal stability. T6r^+ DNA and avian PPLO DNA were denatured and then renatured under optimum conditions in 0·3 M-NaCl plus 0·03 M-Na citrate. Each was then compared to its corresponding native preparations for its thermal stability in the same solvent. From Fig. 9(*a*) and (*b*) it is clearly seen that a large portion of each pair of curves can be superimposed in the T_m region and can be readily accounted for, if one assumes that 70 to 80% of the denatured molecules have renatured when exposed to the optimum renaturation conditions. Moreover, we have recently found (Cordes & Marmur, unpublished observations) that denatured DNA

from the phage α can be renatured to the extent of approximately 90% using the criterion of the extent that the absorbance-temperature curves can be superimposed in the T_m region.

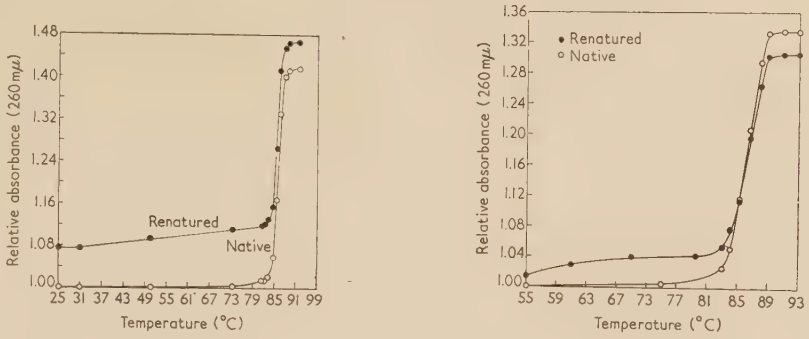


FIG. 9 (*a*). Thermal stability of renatured and native T6r^+ DNA in 0·3 M-NaCl plus 0·03 M-Na citrate. Conditions are described in the text.

(*b*) Thermal stability of renatured and native *Mycoplasma gallisepticum* (avian PPLO 5969) DNA in 0·3 M-NaCl plus 0·03 M-Na citrate. Conditions are described in the text.

4. Discussion

By studying the renaturation of various samples of DNA we have been able to confirm previous notions about the properties of DNA and to advance our understanding of its structure. Thus, the inability to renature DNA from higher plants and animals is predictable because of their much greater heterogeneity and the dependence of renaturation on the concentration of *specific* strands. DNA from relatively homogeneous sources (e.g. bacteriophage) are expected to and indeed do renature at the greater rate and to a higher extent. It remains to be shown whether or not, by fractionating DNA from heterogeneous sources, renaturation of selected, more homogeneous fractions can be attained.

The optimum renaturation temperature is approximately 25°C below T_m. The reason for this may be the result of several factors. An elevated temperature is required for the melting of the weak, intramolecular hydrogen bonds of the fast-cooled, denatured DNA strands as well as to disrupt any mismatched, non-complementary strands. It would be expected that the recognition of complementary strands requires a certain minimum number of base pairs which are large enough in number to resist being melted out at the elevated, optimum temperature. Once the minimum nucleation region has been achieved, the remaining bases of the complementary strands, which are in register, can be imagined to "zip" up to complete the renaturation of the molecule. DNA with a higher guanine *plus* cytosine content would be expected to have a higher optimum renaturation temperature.

Employing optimum conditions, denatured fast-cooled *D. pneumoniae* DNA (with approximately 0·5% residual biological activity) can be renatured to 50 and sometimes 60% of the original transforming activity. The inability to achieve complete renaturation is due to some phosphate-ester backbone breakage of the DNA strands during the denaturation and renaturation (Doty *et al.*, 1960) and perhaps also to depurination (Greer & Zamenhof, 1959) as well as to the presence of some residual

unrenatured molecules. Employing milder denaturation conditions (e.g. formamide) and a more homogeneous DNA source, it would be expected that, under optimum renaturation conditions of DNA concentration, ionic strength and temperature, higher degrees of renaturation will be attained.

The authors wish to thank Mr. C. Schildkraut for his advice and suggestions and Miss D. Lane for participating in the biological experiments.

REFERENCES

Belozersky, A. N. & Spirin, A. S. (1960). In *The Nucleic Acids*, ed. by E. Chargaff & J. N. Davidson, Vol. 3, p. 147. New York: Academic Press.
Brown, G. L. & Brown, A. V. (1958). *Symp. Soc. Exp. Biol.* **12**, 6.
Chargaff, E. (1955). In *The Nucleic Acids*, ed. by E. Chargaff & J. N. Davidson, Vol. 1, p. 521. New York: Academic Press.
Crampton, C. F., Lipshitz, R. & Chargaff, E. (1954). *J. Biol. Chem.* **211**, 125.
Doty, P., Boedtker, H., Fresco, J. R., Haselkorn, R. & Litt, M. (1959*a*). *Proc. Nat. Acad. Sci., Wash.* **45**, 482.
Doty, P., Marmur, J., Eigner, J. & Schildkraut, C. (1960). *Proc. Nat. Acad. Sci., Wash.* **46**, 461.
Doty, P., Marmur, J. & Sueoka, N. (1959*b*). In *Structure and Function of Genetic Elements*, p. 1. Brookhaven Symposia in Biology.
Fox, M. S. & Hotchkiss, R. D. (1957). *Nature*, **177**, 1322.
Greer, S. & Zamenhof, S. (1959). *Fed. Proc.* **18**, 939.
Guild, W. R. (1961). *Fifth Annual Meeting of the Biophysical Society*, FB 9.
Guild, W. R., Morowitz, H. J. & Castro, E. (1960). *Fourth Annual Meeting of the Biophysical Society*, p. 19.
Hershey, A. D. & Burgi, E. (1960). *J. Mol. Biol.* **2**, 143.
Levine, L., Murakami, W. T., Van Vunakis, H. & Grossman, L. (1960). *Proc. Nat. Acad. Sci., Wash.* **46**, 1038.
Litt, M., Marmur, J., Ephrussi-Taylor, H. & Doty, P. (1958). *Proc. Nat. Acad. Sci., Wash.* **44**, 144.
Marmur, J. (1961). *J. Mol. Biol.* **3**, 208.
Marmur, J. & Doty, P. (1959). *Nature*, **183**, 1427.
Marmur, J. & Doty, P. (1961). In preparation.
Marmur, J. & Lane, D. (1960). *Proc. Nat. Acad. Sci., Wash.* **46**, 453.
Marmur, J. & Ts'o, P. O. P. (1961). *Biochim. biophys. Acta*, in the press.
Vendreley, R. (1958). *Ann. Inst. Pasteur*, **94**, 142.

J. Mol. Biol. (1961) **3**, 595–617

The Formation of Hybrid DNA Molecules and their use in Studies of DNA Homologies

Carl L. Schildkraut†, Julius Marmur‡ and Paul Doty

Department of Chemistry, Harvard University, Cambridge 38, Massachusetts, U.S.A.

(*Received 29 March 1961*)

Heavy-isotope-labeled DNA has been used to study strand separation and recombination. The rate at which ^{14}N–^{15}N biologically half-labeled DNA can be made to separate into subunits corresponds very closely to what has been predicted for the rate of unwinding of the strands of a double helix.

The use of a phosphodiesterase from *E. coli* (Lehman, 1960) which selectively attacks single-stranded DNA has made it possible to remove unmatched single chain ends from renatured DNA and reduce the remaining differences between renatured and native DNA.

A mixture of heavy-isotope-labeled and normal bacterial DNA was taken through a heating and annealing cycle, treated with the phosphodiesterase and examined by cesium chloride density-gradient centrifugation. Three bands were observed, corresponding to heavy renatured, hybrid, and light renatured DNA. As would be expected for random pairing of complementary strands, the amount of the hybrid was double that of either the heavy or the light component. It has thus been demonstrated that the strands which unite in renaturation are not the same strands that were united in the native DNA but instead are complementary strands originating in different bacterial cells.

The formation of hybrids has been possible only where the heavy and normal DNA samples have a similar overall base composition. It has also been shown for DNA samples isolated from bacteria of different genera in a case where genetic exchange by conjugation has been demonstrated. The evidence for the parallelism between genetic compatibility and the formation of DNA hybrids *in vitro* has led to the proposal that organisms yielding DNA which forms hybrid molecules are genetically and taxonomically related.

1. Introduction

It has been shown recently that it is possible to separate the two strands of DNA molecules and then to reunite these strands so as to restore to a large degree the original helical structure and biological activity (Marmur & Lane, 1960; Doty, Marmur, Eigner & Schildkraut, 1960). The optimum conditions for bringing about maximum restoration, that is, renaturation, have been presented in the preceding paper (Marmur & Doty, 1961).

The work presented in this paper is aimed at the detailed examination of the process of strand separation and recombination, the formation of hybrid DNA molecules with

† Present address: Department of Biochemistry, Stanford University, Palo Alto, California.

‡ Present address: Graduate Department of Biochemistry, Brandeis University, Waltham, Massachusetts.

each of the strands coming from a different source, and the use of such hybrids in detecting overlap in nucleotide sequences in the two samples. If complementarity is the condition for maximum renaturation then it is proper to enquire to what extent deviations from complementarity are required to interfere with the renaturation. The most modest deviation is probably that which occurs between a point mutant and wild-type DNA. In a following paper (Marmur, Lane & Doty, 1961) this kind of recombination is studied with bacterial transformation as the criterion and it is concluded that such small differences do not interfere with renaturation.

The method of study in the present report was density-gradient centrifugation (Meselson, Stahl & Vinograd, 1957) coupled with the use of heavy-isotope-labeled DNA. This proves to be the best technique for following the interaction between homologous and heterologous DNA. Native DNA samples of different GC (guanine-cytosine) contents can be observed separately when banded in the density gradient (Sueoka, Marmur & Doty, 1959; Rolfe & Meselson, 1959). Two native samples having the same GC content can also be observed and their interaction studied if one of them is labeled with heavy isotopes such as ^{15}N, ^{13}C and deuterium (Meselson & Stahl, 1958; Davern & Meselson, 1960; Marmur & Schildkraut, 1961*a*).

A double helix resulting from the union of a heavy-labeled and non-labeled strand will have an intermediate density. It can be clearly distinguished from the original molecules if they are separated by a sufficient distance in the density gradient. The amount of hybrid molecules formed between different pairs of DNA samples expected to be homologous should vary with the degree of homology.

In order to examine the possibilities of renaturation occurring between DNA from different species that may be genetically related, more quantitative work was necessary to obtain a proper base line for comparison. The removal of non-renatured regions by means of a new phosphodiesterase, specific to single-stranded DNA (Lehman, 1960) provided a means of eliminating artifacts and improving resolution so that hybrid DNA molecules could be clearly and quantitatively resolved. As a result it was possible to examine in some detail the renaturation that was possible between pairs of DNA samples from different sources. This appears to offer a new means of estimating the extent of overlap in DNA sequence between DNA samples from genetically related species.

2. Materials and Methods

(a) *Bacterial strains.* The organisms used in this study together with the base composition and buoyant density of their DNA are listed in Table 1.

(b) *Preparation of DNA.* The isolation procedure of Marmur (1961) was used to extract DNA from cells grown in the exponential phase in Difco brain heart infusion medium. When heavy-isotope-labeled DNA was required the cells were grown in a synthetic medium containing $^{15}NH_4Cl$ as the only nitrogen source and D_2O of greater than 99% purity (Marmur & Schildkraut, 1961*a*).

(c) *Solvents.* In order to eliminate spurious effects due to divalent metal ion contamination, and to inhibit possible attack by nucleases, a chelating agent was always present in the DNA solutions. Citrate ions have proved most convenient in this respect and the standard saline solution (0·15 M-NaCl was used most often) contained 0·015 M Na-citrate. This solvent, standard saline-citrate, will be designated by the abbreviation SSC. Reference will also be made to various multiples of SSC. For example, 2 × SSC is 0·30 M-NaCl and 0·030 M-Na citrate.

(d) *Heating and Annealing.* Solutions containing equal amounts (by weight) of heavy-isotope-labeled and normal DNA in 1·9 × SSC were prepared. In the usual preparation of

renatured DNA, 1 ml. was placed in a 2 ml. glass stoppered volumetric flask, immersed in a bath of boiling water for 10 min and immediately transferred to a bath thermostated at 68°C. After 2 hr the temperature was lowered in approximately 5°C steps at intervals of 15 min until 25°C was reached. Although all results reported here were obtained by this cooling procedure, it has been found that the salt concentration and the rate of cooling from 68°C to room temperature can be varied somewhat without noticeable change in the final results.

TABLE 1

Bacteria from which DNA was isolated for use in studies of hybrid formation

Species	Strain	Source	%GC†	Density (g/cm³)
Bacillus brevis	9999	ATCC‡		1·704
Bacillus macerans	7069	ATCC		1·713
Bacillus megaterium		Univ. of Penn.	38	1·697
Bacillus natto	MB-275	A Demain		1·703
Bacillus subtilis	168	Yale University	42	1·703
Clostridium perfringens	87b	M. Mandel	31	1·691
Diplococcus pneumoniae	R-36A	R. Hotchkiss	39	1·701
Erwinia carotovora	8061	ATCC	54	1·709
Escherichia coli	B	S. Luria	50	1·710
Escherichia coli	44 B	M. Mandel		1·710
Escherichia coli	C-600	R. Appleyard		
Escherichia coli	K12 (W678)	J. Lederberg	50	1·710
Escherichia coli	TAU⁻	S. Cohen	50	1·710
Escherichia coli	W-3110	M. Yarmolinsky		1·710
Escherichia coli	I	A. N. Belozersky	52	
Escherichia coli	II-IV-4	A. N. Belozersky	67	
Escherichia freundii	17	H. Blechman		1·710
Escherichia freundii	5610-52	M. Mandel		
Salmonella arizona	PCl45	Walter Reed Hosp.		1·712
Salmonella ballerup	ETS107	Walter Reed Hosp.		
Salmonella typhimurium	LT-2	M. Demerec	50, 54	1·712
Salmonella typhimurium	ETS9	Walter Reed Hosp.		
Salmonella typhosa	643	Walter Reed Hosp.	53	1·711
Shigella dysenteriae	15	S. Luria	53	1·710

† The GC contents of the DNA samples were all obtained from Belozersky & Spirin (1960).
‡ American Type Culture Collection.

(e) *CsCl.* Optical grade CsCl was obtained from the Maywood Chemical Co., Maywood, N.J. A concentrated stock solution was prepared by dissolving 130 g CsCl in 70 ml. of 0·02 M-tris buffer (2-amino-2-hydroxymethylpropane-1:3-diol) pH 8·5. The final solution was passed through a medium grade sintered glass filter to remove large amounts of solid material that seemed to contaminate the solid CsCl. When solutions in which the DNA had been annealed at low concentrations had to be brought up to the correct density for centrifugation, the solid CsCl was added directly. These solutions were not subsequently filtered.

(f) *Density-gradient centrifugation.* The technique described by Meselson *et al.* (1957) was followed. CsCl was used to bring the density of the DNA solution being examined to values between 1·71 g/cm³ and 1·75 g/cm³, depending on the specific sample involved. This could be done most easily by mixing the DNA solution with the concentrated CsCl stock solution. If the DNA solution was not concentrated enough, as was generally the case, this procedure would diminish the final concentration of the DNA below that required for accurate observation. To avoid this problem it was possible to use solid CsCl to give the proper density. To 1·03 g

CsCl was added 0·80 ml. of the slow cooled DNA solution (still in the same saline citrate solution) and 0·01 ml. of a stock solution of 50 μg/ml. of a DNA sample whose density is well established and can be used as a standard. DNA from *Cl. perfringens* is generally used since its density is the lowest observed thus far. This places it in a position where its band cannot coincide with any heavy-isotope-labeled DNA samples. The final adjustment of density was made by the addition of a small amount of water or solid CsCl. The measurement of the density at various stages of the adjustment is facilitated by the use of the linear relation between refractive index and density (Meselson, 1958, personal communication; Ifft, Voet & Vinograd, 1961)

$$\rho^{25\cdot0°C} = 10\cdot8601\ n_D^{25\cdot0°C} - 13\cdot4974.$$

Approximately 0·75 ml. of the final CsCl solution was placed in a cell containing a plastic (Kel-F) centerpiece and centrifuged in a Spinco model E analytical ultracentrifuge at 44,770 rev/min at 25°C. After 20 hr of centrifugation, ultraviolet absorption photographs were taken on Kodak commercial film. It was clear that equilibrium had been closely approached in 20 hr since there was no difference in calculated densities when photographs taken after 48 hr were used. Moreover, the variances did not decrease by more than 5% between 20 and 48 hr. Tracings were made with a Joyce-Loebl double-beam recording microdensitometer with an effective slit width of 50 microns in the film dimension.

Densities were calculated by using the position of the standard DNA as a reference. The CsCl density gradient was obtained from the data of Ifft *et al.* (1961).

(g) E. Coli *phosphodiesterase.* This enzyme, which preferentially hydrolyses single-stranded DNA, was kindly supplied by Dr. L. Grossman and had been prepared in highly purified form by the method of Lehman (1961, personal communication). The enzymatic hydrolysis was carried out by the following modification (L. Grossman, 1961, personal communication) of the method of Lehman (1960). The annealed sample, usually at a total DNA concentration of 10 μg/ml., was dialysed against two changes of 0·067 M-glycine buffer, pH 9·2. To 0·7 ml. of the dialysed DNA solution was added 1·6 μmoles of $MgCl_2$, 2·4 μmoles of 2-mercaptoethanol, and approximately 25 to 75 units of crystallized *E. coli* phosphodiesterase. The final mixture (about 0·8 ml.) was incubated at 37°C for 3 hr. Solid CsCl and the standard DNA were then added and the solution was made up to the proper density.

3. Strand Separation

In order to form hybrid molecules by renaturation it is first necessary to examine in some detail the conditions necessary for strand separation. A very useful material for these studies has been the ^{14}N-^{15}N-half-labeled or biological "hybrid" DNA molecules first described by Meselson & Stahl (1958). DNA having one strand ^{15}N-labeled and the other strand normal was isolated from *E. coli* B according to the method of Marmur (1961). It was found that the biologically formed "hybrid" DNA banded at a density of 1·717 g/cm^3 in the CsCl gradient, exactly between that of fully ^{15}N-labeled and normal DNA from *E. coli*. The band profile is shown in the top tracing of Fig. 1.

(a) *Methods for inducing strand separation*

In terms of the Watson-Crick model (1953) for DNA, the dissociation of the two strands obviously requires first the rupture of the hydrogen bonds uniting the base pairs and second the uncoiling and diffusing apart of the two strands. The rupture of the hydrogen bonds has been shown to occur in a relatively narrow temperature range at elevated temperatures but the diffusion apart of the strands did not automatically follow (Rice & Doty, 1957). Later it was found that lower ionic strengths and more dilute solutions did permit the diffusing apart to occur (Doty *et al.*, 1960;

Eigner & Doty, 1961) and prevented the non-specific re-association from occurring at lower temperatures where numerous hydrogen bonds re-form.

Meanwhile it had been found by Meselson & Stahl (1958) that the DNAs containing ^{14}N and ^{15}N subunits, each of which replicate in a conservative manner, produced two bands on heating in CsCl for 30 min. Thus they demonstrated that the two molecular subunits dissociated on heating. Meselson & Stahl left the question of the molecular structure of the subunits open to further investigation. This prompted us to isolate DNA from cells grown according to the procedure described by Meselson & Stahl (1958), and to see how the hybrid DNA reacted to the various methods that produce strand separation. The evidence that these methods do produce strand separation can be found in the papers to which we will refer. We are concerned here only with their effect on the biological hybrids and the detailed investigation of this effect in order to learn more about strand separation.

First we sought to reproduce the Meselson-Stahl results. They had observed the two-band pattern after heating a cell lysate in CsCl for 30 min. We found a similar pattern was produced by heating a sample of our purified biological hybrid at 100°C for only 10 min at a concentration of 20 μg/ml. in SSC. The tracings of the bands before and after heat treatment can be seen at the top and bottom frames of Fig. 1. The apparent density of the native, hybrid DNA is 1·717 g/cm^3; that of the two denatured bands is greater by 0·007 and 0·023 g/cm^3 respectively. The average of these is, of course, greater than that of the native hybrid DNA by the 0·015 g/cm^3 characteristic of denatured DNA. As a check, it was demonstrated that the same two-band pattern results when ^{14}N and ^{15}N DNA samples are separately denatured and then mixed.

Three other methods which have recently been shown to bring about strand separation have also caused the disappearance of the native hybrid band and the formation of two heavier bands. This additional information, combined with the demonstration of Doty *et al.* (1960) that the above heating procedure produces strand separation in bacterial DNA samples, now leads us to conclude that Meselson & Stahl were indeed observing the separation of single strands. We shall return to this point in the discussion, and now proceed to summarize the experimental observations on the biological hybrids.

By the addition of 8 M-urea the temperature for thermal denaturation can be reduced by nearly 20°C (Rice & Doty, 1957). Consequently, by using 0·01 M-salt and 8 M-urea it was found possible to bring about strand separation by heating DNA only to 65°C (Eigner & Doty, 1961). When this procedure was applied to the biological hybrid DNA, two bands formed in the CsCl density gradient, exactly as was the case after heating at 100°C for 30 min in CsCl or for 10 min in SSC.

Formamide, being a more potent hydrogen bonding agent, and being miscible with water in all proportions, offers a substantial improvement over urea. Marmur & Ts'o (1961) have shown that in 95% formamide strand separation readily occurs at room temperature. Again, when this is applied to the biological hybrids the typical two-band pattern is produced.

Acid and base titration, by virtue of their ability to disrupt hydrogen bonding, offer another route to strand separation. It has now been shown (Cox, Marmur & Doty, 1961) that strand separation occurs when the pH is lowered to 2·5 or raised to 12·0. These two pH values lie just outside the region in which the hypochromic shift (at 260 mμ) indicates that the helix-coil transition is complete. At pH values on

the other side of the hypochromic shift, strand separation has not occurred. When the biological hybrids are exposed to conditions in which the pH is below 2·5 or above 12·0, two bands again result in the CsCl gradient.

Thus it is seen that several different procedures lead to the formation of the same two-band pattern in the CsCl density gradient, and that all of them are a consequence of the prior breaking of hydrogen bonds uniting the two DNA strands. While the present work is not aimed at establishing how many subunits make up DNA molecules, it does appear that the extension to the biological hybrid DNA of the consistent finding that strand separation occurs whenever hydrogen bonds are cooperatively broken, shows that only one type of bonding is uniting the subunits. Since this can only be identified with that assumed in the Watson-Crick structure, there seems to be no alternative to equating the subunits, observed as two bands, with the two strands in the native DNA molecule.

(b) *Temperature dependence of strand separation*

With the demonstration that the two-band pattern of the biological hybrid DNA is indicative of strand separation we now attempt to use this material to obtain further details about the process of strand separation. The studies of Marmur & Doty (1959) demonstrated that it should be possible to melt out the molecules having a high AT (adenine-thymine) content leaving GC (guanine-cytosine)-rich molecules undenatured. This can be observed to occur with the biological hybrids by exposing them to temperatures only slightly above the temperature, T_m, of the midpoint of the absorbance rise. T_m is the same for all *E. coli* DNA samples, including the biological hybrids (Marmur & Doty, 1961). The band profile obtained after heating the latter at T_m for 20 min is shown as the second tracing in Fig. 1. Since there is no difference from the profile of the native sample, it is concluded that all the molecules remain in the helical form. To melt selectively the AT-rich molecules, it is necessary to expose them to temperatures a few degrees above T_m. The results of doing this are shown as the third tracing of Fig. 1, where it is seen that approximately 25% of the molecules undergo strand separation. The molecules showing the greater resistance to heat denaturation are seen to have a higher density, in agreement with the hypothesis that they should be GC-rich.

(c) *Kinetics of strand separation*

Since a small but detectable thermal depolymerization always accompanies strand separation it is desirable to know the minimum time at the elevated temperature required to obtain complete separation, and thus avoid this undesirable side effect. In addition, since a few theoretical calculations of the rate of uncoiling of the DNA molecule have been reported (Kuhn, 1957; Longuet-Higgins & Zimm, 1960), some comparison with experimental findings seems to be in order. The rate of separation into subunits can be observed experimentally by following the disappearance of the hybrid band or the appearance of the two heavier bands. The results presented in Plate I and Fig. 2 show that the complete disappearance of material of density 1·717 g/cm^3 takes slightly over 60 sec at 100°C. This is very close indeed to the time calculated by Kuhn (1957) for the unwinding of a double helix of comparable length in which no bonds are holding the strands together. As is evident from the absorbance-temperature profile, the rise in absorbance is complete at 95°C. When the temperature of a DNA solution

PLATE I. Ultraviolet absorption photographs of ^{14}N-^{15}N-labeled DNA showing different stages of the thermally induced separation into subunits. The band at the far right has been used as a standard and is DNA isolated from *D. pneumoniae*. The other band in the top photograph is the biologically formed hybrid DNA. The second photograph shows the stability of the hybrid to a 20 min exposure at 93·8°C. At 100°C the number of molecules separating increases rapidly with time of exposure as shown in the next 3 photographs. The samples were heated in SSC at 20 μg/ml. for 30 sec, 1 min and 10 min, respectively.

PLATE II. Electron micrograph of renatured *D. pneumoniae* DNA showing an unusually large concentration of renatured molecules with circular regions at one or both ends. Magnification $\times$100,000.

[*To face page* 600

is raised to 100°C few hydrogen bonds should exist between strands and Kuhn's model seems applicable to such a molecule.

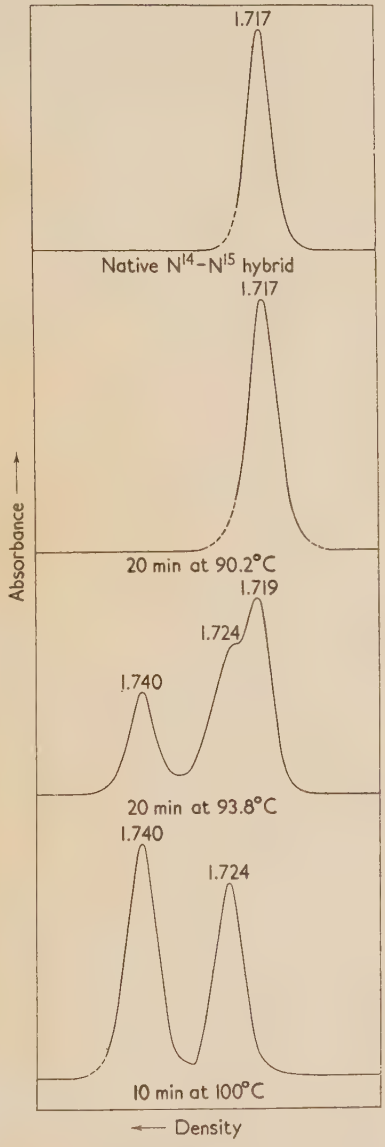


FIG. 1. The temperature dependence of strand separation. Microdensitometer tracings of ultraviolet absorption photographs of samples equilibrated in a CsCl density gradient. Each sample was heated in SSC at a concentration of 20 μg/ml.

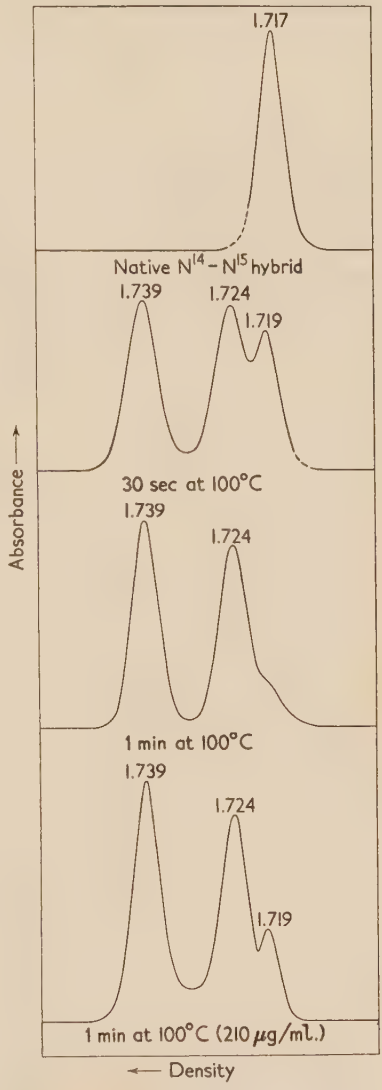


FIG. 2. Kinetics of strand separation. The results shown in the second and third tracings were obtained by heating at a concentration of 20 μg/ml. in SSC.

(d) *Dependence of the time of strand separation on viscosity*

In the theoretical considerations just mentioned in connection with the kinetics of strand separation, the variation of the rate of unwinding with the viscosity of the

medium is also considered (Longuet-Higgins & Zimm, 1960). An equation is derived which predicts that strand separation will take longer as the viscosity of the medium is increased. In the tracing shown at the bottom of Fig. 2 the time necessary for separation into subunits is shown to have increased with increased DNA concentration. Similar results have also been obtained by using concentrated sucrose solutions to produce a medium of higher viscosity (Schildkraut, Wierzchowski & Doty, 1961).

(e) *DNA from bacteriophage ΦX174*

The results of the studies with the biological hybrids indicate that the most convenient way to obtain single-stranded DNA is to heat for 10 min in SSC at a concentration of from 10 to 20 μg/ml. and quickly cool. Since one kind of DNA, that from bacteriophage ΦX174†, is known to be single-stranded (Sinsheimer, 1959), it is expected that our thermal treatment would produce no density change in this material. This has, indeed, been found to be the case.

4. Renaturation

The general features of the specific recombination of single DNA strands to form the native helical structure have been described in the first publications (Marmur & Lane, 1960; Doty *et al.*, 1960), and methods of optimizing conditions to ensure the maximum renaturation were described in the preceding paper (Marmur & Doty, 1961). Here we have only one point to add to this: it has to do with the reduction or elimination of the remaining differences between renatured and native DNA. In the earlier work it was evident that renaturation was not complete. For example, the density had returned only about 75% of the way from the denatured to the native density. It was thought that this was due to the inequality in length of the recombined strands. DNA molecules are apt to be broken during isolation and the single strands suffer some hydrolysis during the thermal treatment. Using the measurements on rate of bond scission by Eigner, Boedtker & Michaels (1961), we would estimate that strands originally in molecules of 10,000,000 molecular weight would undergo about three scissions each as a result of a typical heat treatment: 10 min at 100°C and 100 min at 68°C. This would mean that two strands meeting in a complementary region may differ in length by as much as one quarter to one half, on the average. If they renatured completely, there would still remain 20 to 33% of the weight unrenatured and the incomplete recovery of the characteristics of the native DNA would be explained.

In the course of electron microscope studies with Professor C. E. Hall we frequently saw renatured molecules with circular regions at one or both ends that would be consistent with the protrusion of a single chain end beyond the re-formed helical region just described. One electron micrograph showing an unusually large concentration of these in a single photograph is shown in Plate II. While this interpretation lacks proof, the observation is nevertheless strikingly similar to what was expected from the point of view outlined.

From this evidence it appeared that in order to improve the extent of renaturation the unmatched single chain ends would have to be removed. The possibility of doing this suddenly became available with the discovery by Lehman (1960) of a phosphodiesterase from *E. coli* that selectively attacked single-stranded DNA. This was tested

† This material was kindly provided by Professor Robert L. Sinsheimer.

on renatured DNA from *B. subtilis* as shown in Fig. 3. The denatured DNA showed the characteristic increase in density. Upon renaturation at a concentration low enough to ensure that some denatured DNA remained, the result shown in the third frame was obtained. The density of the renatured DNA is, as expected, 0·004 g/cm^3 higher than that of the native DNA. After treatment of this sample with the phosphodiesterase it is seen, in the bottom frame, that the denatured shoulder has been

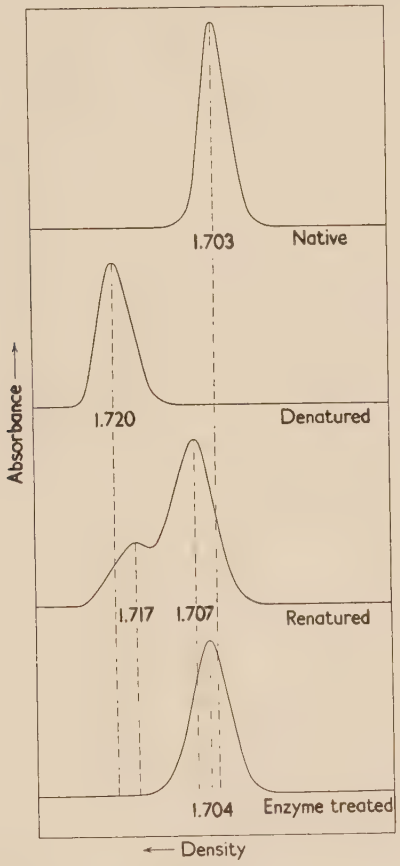


FIG. 3. Renaturation of *B. subtilis* DNA. When a solution of native DNA (band profile shown in top tracing) is heated for 10 min at 10 μg/ml. at 100°C in 1·9 × SSC and quickly cooled, the density increases 0·017 g/cm^3 (second tracing). A portion of this solution was annealed and, as is evident from the third tracing, about 80% of the DNA renatures. This concentration (10 μg/ml.) was chosen so that some denatured material would still be present. Treatment with the *E. coli* phosphodiesterase causes the complete disappearance of the denatured band and a decrease in the buoyant density of the renatured band (bottom tracing).

completely removed and the density of the renatured band has been reduced to within 0·001 g/cm^3 of the native DNA density. Thus the possibility of removing the denatured unmatched ends from renatured molecules has been demonstrated. As will be seen, this procedure greatly improves the resolution of bands encountered in the study of artificially produced hybrid molecules.

It may be of interest to mention at this point our failure thus far to produce renaturation of DNA from bacterial sources without thermal treatment of the type described here and in the preceding papers. We have employed urea and moderate

temperatures, formamide, and low pH to produce strand separation; and then, by gradual withdrawal of the hydrogen-bond breaking agent attempted to renature the strands. Such attempts have thus far failed. While our experiments have not been exhaustive, it appears quite possible that renaturation cannot be achieved in this way. If so, a likely explanation would be that at room temperature the decreased Brownian motion of the chain segments is no longer sufficient to provide the mobility required for the exploration necessary to create nuclei, that is, for complementary regions to find each other rather than become frozen in mismatched pairings.

In connection with renaturation, we might return just briefly to the case of ΦX DNA and mention that the density does not decrease but remains exactly 1·723 g/cm^3 after heating and annealing. This is as should be expected for single-stranded material whose complementary strands are not present during the annealing process.

5. Hybrid DNA Molecules produced by Renaturation

We now turn to the problem of demonstrating that the strands which unite in renaturation are not the same strands that were united in the native DNA but instead are complementary strands originating in different cells. This requires that renaturation be studied in a solution of two homologous DNA samples, one of which carries a distinctive label. With density gradient ultracentrifugation offering such good resolution it was natural to turn to the introduction of heavy atoms or to heavy isotope substitution. If the pairing of DNA strands is restrained only by the condition of complementarity one would expect that renaturation would lead to three bands, one heavy, one light and one intermediate, corresponding in density to a hybrid composed of one normal and one heavy strand. Moreover, the amount of the intermediate would be expected to be double that of either the heavy or light component provided that equal amounts of the two DNA samples had been mixed originally.

If this situation is found, then it will be of interest to attempt to form hybrid molecules from other pairs of DNA samples in which some differences exist. In order to have sufficient resolution to exploit this means of searching for such effects it is necessary to have a considerable density difference between the two samples. Practical considerations indicate that it should be at least 0·040 g/cm^3. Thus the use of ^{15}N label in one sample, such as first used by Meselson & Stahl (1958), is not sufficient since this gives a separation of only about 0·015 g/cm^3.

(a) *Preliminary experiments*

Our first attempts to produce DNA with substantially higher density were with 5-bromouracil substitution for thymine inasmuch as this had been accomplished in both bacteria and bacteriophage (Dunn & Smith, 1954; Zamenhof & Griboff, 1954). A sample of such DNA from *B. subtilis* (Ephrati-Elizur & Zamenhof, 1959) was kindly supplied to us by Professor W. Szybalski. Unfortunately, when this sample was carried through the heating and annealing cycle it displayed a quite broad band with little indication of renaturation. The molecular weight had evidently been substantially reduced either by enzymatic attack during isolation or as a result of the greater heat sensitivity of this substituted DNA.

Nevertheless this material was mixed with high molecular weight, normal *B. subtilis* DNA, and put through the heating and annealing cycle. The banding of the mixture before and after heating is shown in Fig. 4. It is seen that the thermal treatment has

produced a broad band without any resolution of the expected hybrid. Indeed the result is such as to indicate that the smaller heavy-labeled chains have been bound in a random manner to the larger, normal density chains producing a complete spectrum of densities which hides what little renaturation may have occurred.

Initial attempts to study 5-bromodeoxyuridine-labeled DNA from T4r^+ bacteriophage led to similar disappointing results.

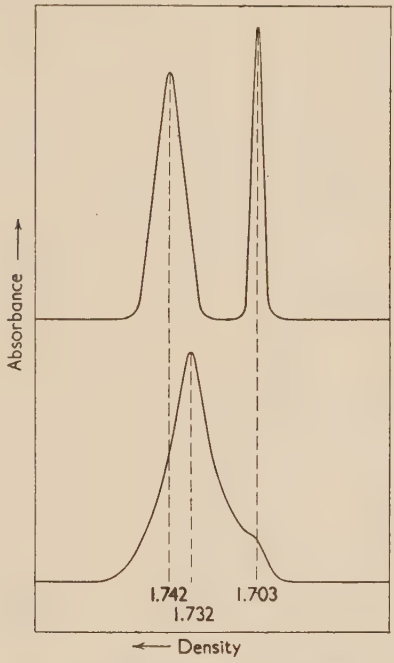


FIG. 4. Hybrid formation using 5-bromodeoxyuridine-labeled DNA. The upper tracing shows native normal and 5-bromodeoxyuridine-labeled DNA isolated from *B. subtilis*. The properties of the latter DNA have been described in detail elsewhere (Szybalski, Opara-Kubinska, Lorkiewicz, Ephrati-Elizur & Zamenhof, 1960). Portions of these two samples were heated together in 1·9 × SSC at a concentration of 20 μg/ml. each for 5 min at 100°C. The solution was then allowed to cool to room temperature over a period of about 5 hr and CsCl was added to obtain the proper density. The lower tracing suggests that the expected renatured species are present in solution, but the broadening of bands due to degradation and aggregation obscures the results.

(b) ^{15}N-*deuterated DNA*

Under these circumstances we turned to a combination of ^{15}N and deuterium to provide the density increase. After considerable manipulation this combined labeling has been shown to be successful in producing DNA with a density increase of about 0·040 g/cm^3 and the details have now been published elsewhere (Marmur & Schildkraut, 1961*a*).

We were, therefore, in a position to proceed with the experiment that would demonstrate whether or not hybrid DNA molecules form from random mating of complementary strands. The steps are shown for *B. subtilis* DNA in Fig. 5. In the first two frames the density profiles of the normal and labeled native DNA samples are shown. These were then mixed so that the concentration of each was 5 μg/ml. in 1·9 × SSC, and the heating and annealing cycle was carried out. At this concentration only partial renaturation is expected. So low a concentration was chosen in order to

minimize the association of renatured DNA molecules since this would lead to an unnecessary smearing of the pattern. Consequently, at this concentration, one would expect to find 5 species of different density: heavy denatured DNA, heavy renatured DNA, hybrid DNA, light denatured DNA and light renatured DNA. The result, seen in the third frame of Fig. 5, is of precisely this character.

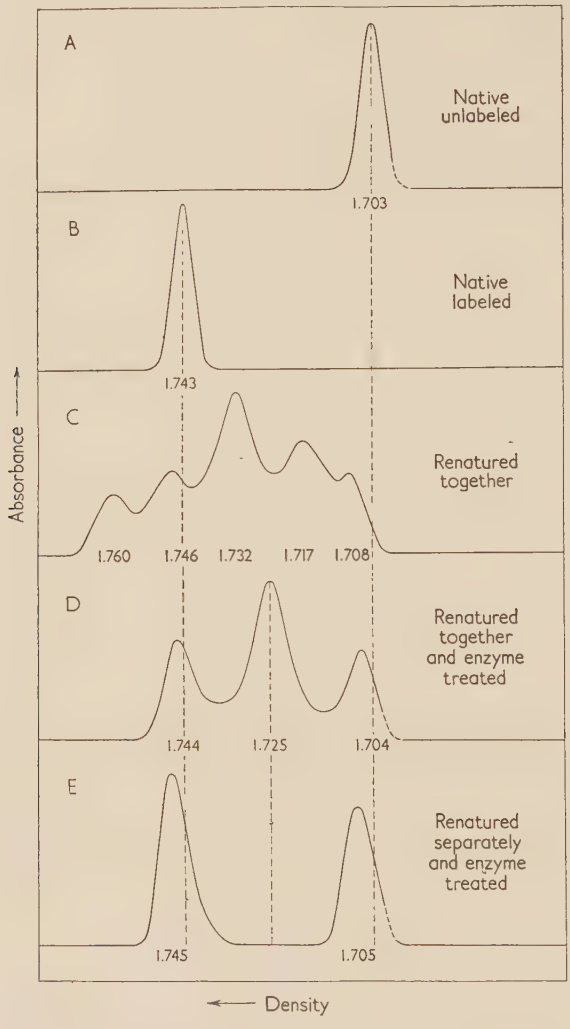


FIG. 5. The effect of *E. coli* phosphodiesterase on a heated and annealed mixture of heavy-labeled and normal *B. subtilis* DNA.

(A) and (B) Native samples.

(C) Mixed so that the final concentration of each was 5 μg/ml. and then heated and annealed.

(D) The sample shown in (C) was dialysed against 0·067 M-glycine buffer and incubated with the *E. coli* phosphodiesterase.

(E) Heated and annealed separately, treated with the phosphodiesterase and then mixed.

The application of the phosphodiesterase previously described would be expected to remove the first and fourth bands as well as shift the density of the remaining species back toward the native values. The band profile after this treatment is shown in the fourth frame and it is seen that the expected changes had been brought about.

As a control, the two native DNA samples were carried through the identical operations described above and then mixed and banded. The result is shown in the last frame where it is seen that the densities match within 0·001 g/cm^3 that of the outer bands of the frame above and that no hybrid is evident.

Similar experiments were carried out on *E. coli* DNA: the band profiles corresponding to frames C and D of Fig. 5 are shown in Fig. 6. The result is seen to be the same.

It is to be noted that the area under the hybrid band in these two experiments is about equal in area to that of the two outer bands.† Thus the conclusion can be drawn that the earlier and much more primitive experiment along these lines (Doty *et al.*, 1960) has been substantiated and the formation of hybrid DNA molecules by random pairing of complementary strands can be taken as proved.

It is useful to note in passing that the phosphodiesterase treatment had one other effect not evident in the band profiles. Without enzyme treatment part of the hybrid band would begin to form much earlier than the other, indicating high molecular weight material suggestive of some aggregation despite our efforts to avoid it. However, after enzyme treatment all three bands appear at the same rate. Thus, the enzymatic treatment not only improves the resolution by removing denatured material but appears to break up aggregated renatured DNA molecules as well.

(c) *Further investigation of 5-bromodeoxyuridine-substituted DNA*

With the successful conclusion of the above work, we have returned again to the 5-bromodeoxyuridine-labeled bacteriophage DNA. With the guide that the above work provides it has been possible to overcome some of the problems that appear to be associated with the greater inherent instability of this material. As a consequence results nearly as satisfactory as those shown in Fig. 5 and 6 are being obtained at present. This system has been used to study DNA homologies among the T-even bacteriophages (Schildkraut, Marmur, Wierzchowski, Green & Doty, 1961).

(d) *Kinetics of hybrid formation*

As a further check on the above interpretation and in order to show further details of the process of hybrid formation, samples were withdrawn from the annealing mixtures, held at 68°C, during renaturation. These were quickly cooled, in order to "freeze" the distribution present and band profiles were obtained in the ultracentrifuge. A number of profiles are shown in Fig. 7. From these, the progress of the renaturation can be followed and it is seen that it reaches about 50% completion in the first hour.

6. Hybrid DNA Formation among Various Strains of *E. coli*

It is to be expected from their close taxonomic, physiological and genetic relationships that all strains identified as *E. coli* should yield DNA which will form the 5-band pattern upon the renaturation discussed in the previous section. Hybrid formation in a number of pairs has been examined. The results are summarized in Plate III where the ultraviolet absorption photographs are reproduced. It is seen that the 5-band pattern is produced in all but the one case where sequence homology did not exist.

† The linearity of response of our optical system is not yet such as to justify a precise measurement of the area under these curves.

This DNA isolated from *E. coli* 11-IV-4, an alkali-producing form, had a GC content of 67%, which is far from that characteristic of *E. coli* DNA. Belozersky (1957) has reported that these cells were produced by an alteration of the properties of *E. coli* CM caused by growing it together with Breslau bacteria No. 70 killed by heat.

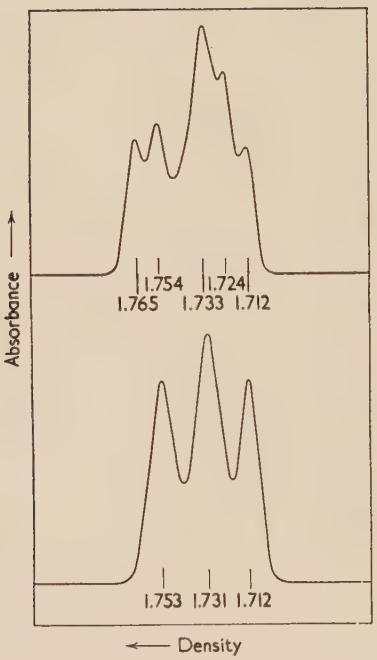


FIG. 6. The effect of *E. coli* phosphodiesterase on a heated and annealed mixture of heavy-labeled DNA from *E. coli* B and normal DNA from *E. coli* K12. The concentration of each sample was 5 μg/ml. each during the heating and annealing cycle.

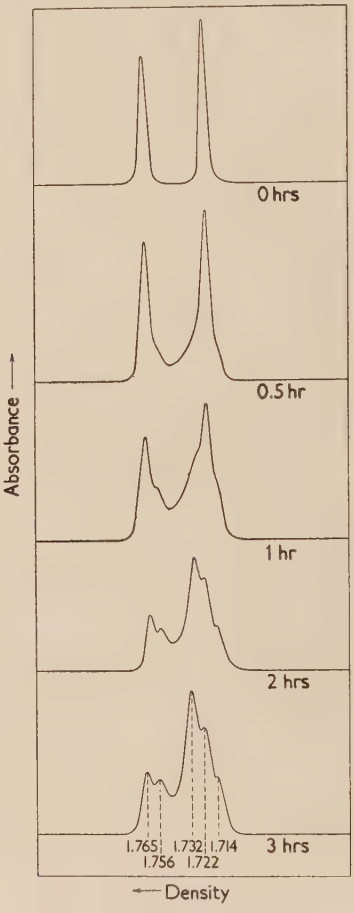


FIG. 7. Kinetics of hybrid formation. Heavy-labeled and unlabeled DNA from *E. coli* B were heated together at 100°C for 10 min under the usual conditions of 5 μg/ml. each and in 1·9 × SSC. The mixture was placed at 68°C and quickly cooled portions were centrifuged after the addition of CsCl. The tracings are shown above. These samples have, of course, not been treated with the phosphodiesterase.

The failure to obtain a hybrid band confirms the gross difference from *E. coli* DNA, but of course does not help to eliminate the argument that these special cells may have been a contaminant rather than a variant. As seen in the photograph the location of the peaks and their intensities is variable. For this reason three of the samples were treated with phosphodiesterase. When this was done the patterns

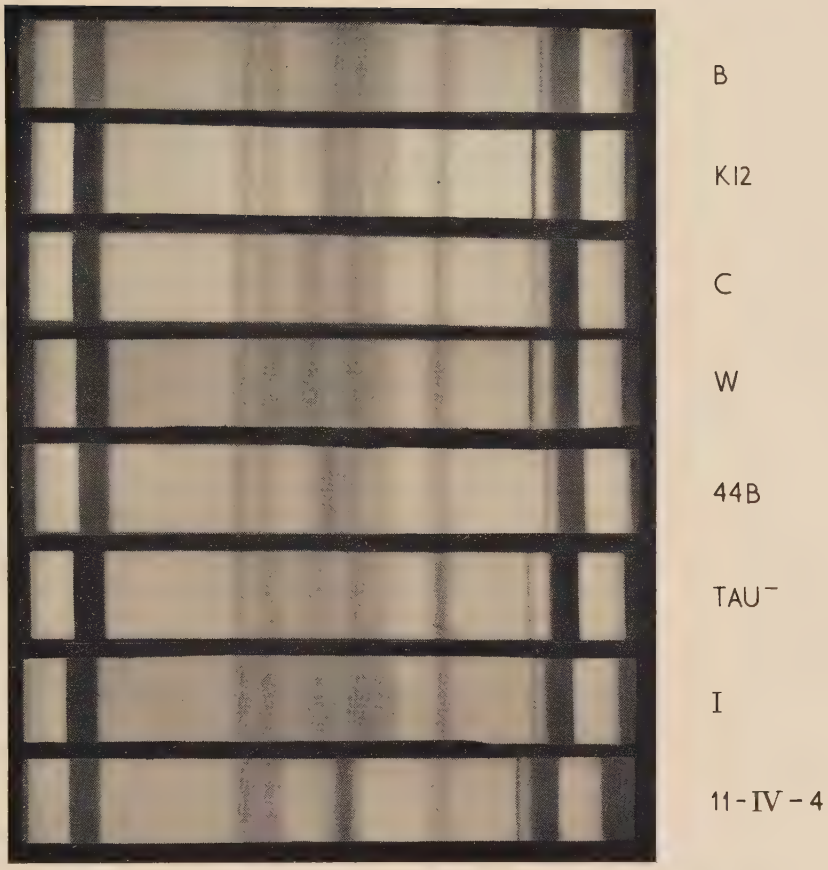


PLATE III. Hybrid formation between *E. coli* B and other *E. coli* strains. *E. coli* B DNA, labeled with ^{15}N and deuterium, was mixed with DNA from each of the strains listed above and heated and annealed in separate experiments. The concentrations were 5 μg/ml. each, and the other conditions were as described in the section on methods. Each of 8 different ultracentrifuge runs is represented above by a typical ultraviolet absorption photograph. Six DNA bands appear in all but the last example. The photographs have been lined up according to the position of the standard band at the far right, which is DNA from *Cl. perfringens*.

[*To face page* 608

gave three bands and showed greatly reduced differentiation. We are, therefore, now in a position to decide if there are any small reproducible differences in the renaturation of various *E. coli* pairs that may indicate small differences in sequence, but the effect is at most quite small. We can conclude that the homology is essentially complete.

7. The Aggregation of Renatured DNA

In exploring further the extent to which hybrids can form between somewhat different DNA strands we encountered more examples of the way in which renatured DNA molecules could aggregate so as to indicate hybrid formation even though it was indeed an artifact. While these difficulties were always removed by phosphodiesterase treatment it is of interest to record a few examples in order to indicate the nature of this process and to emphasize the need for phosphodiesterase treatment.

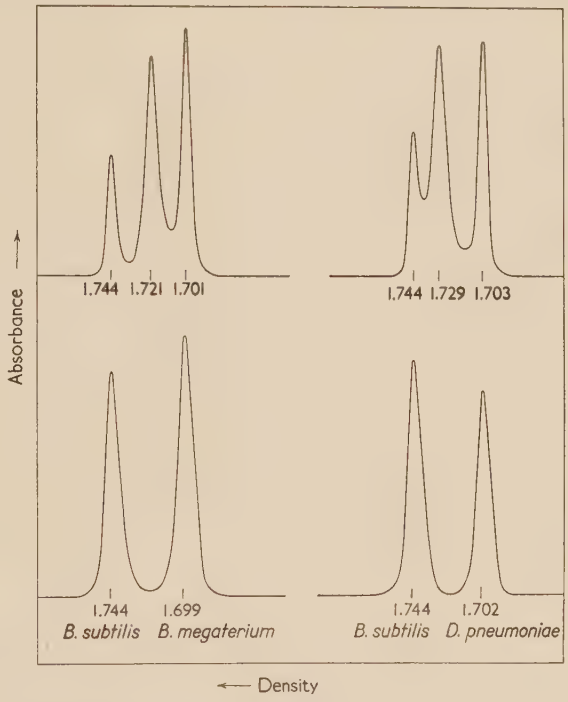


FIG. 8. Non-specific aggregation. ^{15}N-deuterated DNA from *B. subtilis* was heated with either DNA from *B. megaterium* or *D. pneumoniae* at a concentration of 50 μg/ml. each in 1·9 × SSC. The mixtures were annealed under the usual conditions. Portions were centrifuged in CsCl before (upper tracings) and after (lower tracings) treatment with the *E. coli* phosphodiesterase.

A number of experiments were carried out using the heavy-labeled *B. subtilis* DNA and various normal DNA samples at concentrations of about 50 μg/ml. where no denatured strands remain after renaturation. (This is in contrast to the standard procedure which involves tenfold lower concentrations.) Results are shown in Fig. 8 for renatured mixtures of *B. subtilis* DNA with either *B. megaterium* or *D. pneumoniae* DNA. The band profiles before enzyme treatment are shown at the top. Here we see three well-resolved band patterns highly suggestive of hybrid formation. However, the enzyme treatment completely eliminates the central band and increases the amount

of material in the outer bands. Thus the center band must have been the result of homogeneous renatured DNA molecules being held together by the association of unpaired, protruding chains. The rather remarkable feature is that the aggregation must consist mostly of two such renatured molecules, one of each type, since otherwise a whole range of densities would be displayed.

It was also possible to form aggregates between *B. subtilis* and *B. brevis* DNA by heating and annealing at a concentration of 40 μg/ml. each, and between *B. subtilis* and *B. macerans* DNA by heating and annealing at 50 μg/ml. each. The GC content of *B. subtilis* DNA is similar to that of *B. brevis* while that of *B. macerans* is somewhat higher. In this case the density of the aggregates was considerably higher than the average of the two renatured "parent" samples. This was true for practically every case of aggregation except two, one shown in Fig. 8 and the other in Fig. 13. When *B. subtilis* and calf thymus DNA were heated together and annealed at a total concentration of 100 μg/ml., the heavy labeled *B. subtilis* DNA reformed completely while the calf thymus DNA maintained the denatured density. No intermediate band was visible.

As a final example we report the rather puzzling case of aggregation occurring upon mixing renatured solutions at room temperature. Heavy and light *B. subtilis* DNA at a concentration of 10 μg/ml. were heated and annealed separately; the band profiles are shown in the first 4 frames of Fig. 9. The two renatured solutions (C and D) were then mixed at room temperature and examined in the density gradient. The result is shown in Fig. 9 (E); it is seen that a central band has formed, apparently at the expense of the renatured molecules. Thus it would appear that unpaired chain ends which cannot be satisfied in the separated solutions can pair in the mixture. However, this artifact too was removed by phosphodiesterase treatment, as shown in the final frame of Fig. 9. It is important to note that this aggregation of renatured molecules at room temperature occurred between homologous molecules, not molecules from two different sources.

Thus the formation of bands of intermediate density that have the appearance of true hybrid DNA molecules has been clarified, and the importance of eliminating such false bands with phosphodiesterase treatment has been emphasized.

8. Interspecies Hybridization of DNA

We are now in a position to examine the possibility of forming DNA hybrids between strands of DNA coming from different bacterial species. By heating and annealing each of several DNA samples with the heavy-labeled *B. subtilis* DNA, it became evident, even without recourse to the enzyme treatment, that hybrid DNA molecules did not form. Such results are shown in Fig. 10. These and other results could be summarized by stating that hybrid DNA formation was not observed, even at double the usual DNA concentration, if the two DNA samples had significantly different base-compositions.

With the search for hybrid formation narrowed to DNA samples having essentially the same composition, it was natural to look first at DNA from two organisms of the same species which were known to be genetically related by virtue of their ability to undergo transformation with each other's DNA. This is true for *B. subtilis* and *B. natto* (Marmur, Seaman & Levine, 1961). The results for heating, annealing and enzyme treating this pair of DNA samples are shown in Fig. 11. It is seen that substantial

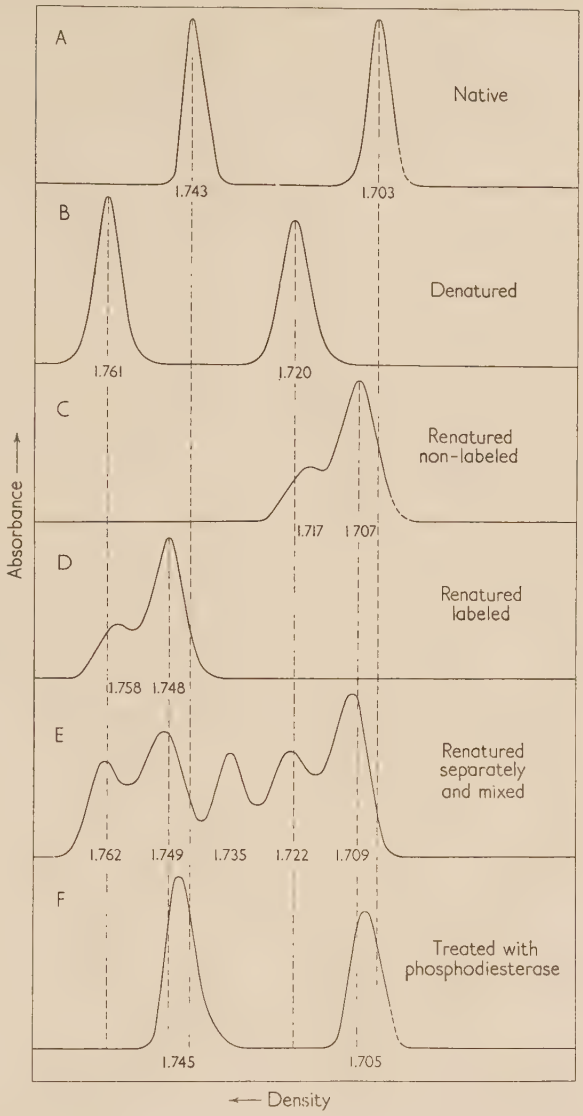


FIG. 9. Control experiment for hybrid formation with *B. subtilis* DNA. A stock solution in 1·9 × SSC of normal *B. subtilis* DNA was prepared so that the concentration was 10 μg/ml. The same was done for an ^{15}N-deuterated sample. After portions of these solutions received the following treatments, they were made up to the proper density with CsCl and centrifuged. The tracings are shown above.

(A) Mixed while still native.

(B) Heated separately for 10 min at 100°C, quickly cooled and mixed.

(C) and (D) Heated separately, annealed under identical conditions and banded separately.

(E) Mixed immediately after annealing while still in the 1·9 × SSC solvent. If the mixture was dialysed against a lower salt concentration (0·067 M-glycine buffer) before addition of CsCl or if each annealed sample was dialysed against 0·067 M-glycine buffer before mixing, 5 bands resulted.

(F) Mixed, dialysed against 0·067 M-glycine buffer and treated with the *E. coli* phosphodiesterase.

The same picture also resulted when portions of the annealed samples shown in B and C were separately treated with the enzyme and then mixed.

hybrid formation did occur although it may not have been as much as expected from random pairing of complementary strands since the three bands appear to have about equal areas.

For the next experiment we chose three members of the family Enterobacteriaceae which have similar base compositions (Lee, Wahl & Barbu, 1956) and which are genetically related. *E. coli* K12 shows a high degree of genetic exchange by conjugation and transduction with *E. coli* B and *Sh. dysenteriae* (Lennox, 1955; Luria & Burrous, 1957; Luria, Adams & Ting, 1960) whereas *Salm. typhimurium* mates well with K12 but is transduced only to a very limited extent, if at all (Zinder, 1960).

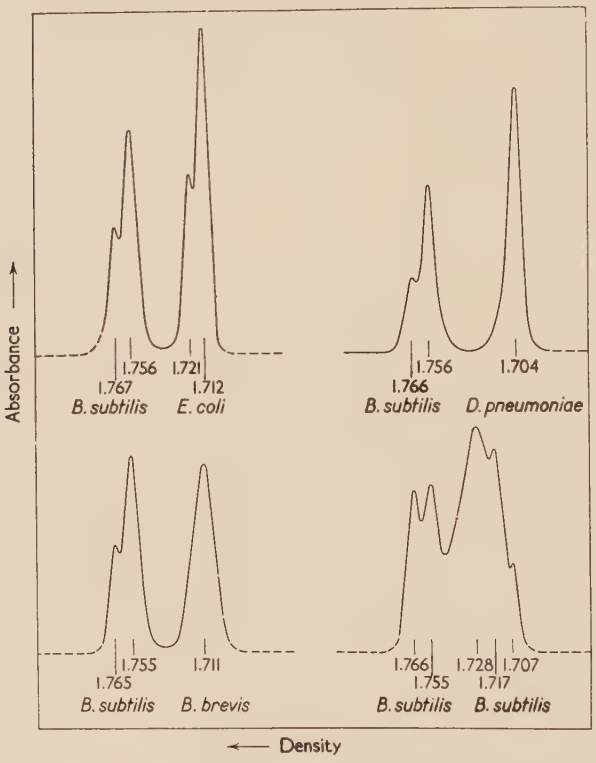


Fig. 10. The specific requirement for hybrid formation using heavy-labeled *B. subtilis* DNA. The three non-homologous samples were heated and annealed at a concentration of 10 μg/ml. each, under the same conditions already described. In the case where both samples were *B. subtilis* DNA, the concentrations were 5 μg/ml. each. Neither the labeled nor the unlabeled *B. subtilis* preparations were the same ones used to obtain the results shown in Fig. 5.

In this context, attempts were begun to prepare hybrid DNA molecules between ^{15}N-deuterated *E. coli* DNA and normal *Shigella* or *Salmonella* DNA. Unlabeled DNA isolated from *Sh. dysenteriae* was substituted for the unlabeled *E. coli* DNA in the procedure discussed in Section 5, and the heated and annealed mixture was treated with *E. coli* phosphodiesterase. Three bands were observed as shown in Fig. 12. The hybrid band does not contain twice as much DNA as either of the uniformly labeled renatured bands, as was observed in most previous examples. This indicates that not every *Sh. dysenteriae* DNA molecule is homologous to a corresponding *E. coli* DNA molecule. Genetic evidence also supports this partial degree of homology (Luria & Burrous, 1957).

Similar studies were carried out with DNA isolated from different strains and species of *Salmonella*. Heating and annealing was carried out with DNA isolated from each *Salmonella* organism listed in Table 1. The results for all samples were similar to those shown in Fig. 13 for the *Salm. typhimurium* and the heavy-labeled *E. coli* B DNA. No hybrid appears at 10 μg/ml. (top tracing) but an intermediate band does appear at 20 μg/ml. (middle tracing). It is, however, removed by enzyme treatment. The experiments were also repeated using heavy-labeled DNA from *E. coli* K12 since this strain has been used most successfully in the studies of genetic recombination. So far,

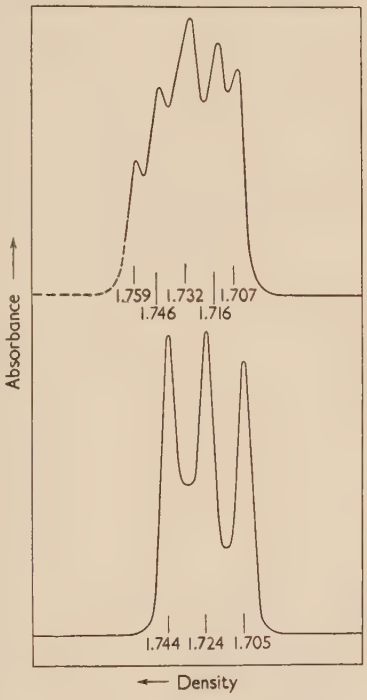


FIG. 11. Interspecies hybrid formation with *B. natto* DNA. Heavy-labeled *B. subtilis* DNA and unlabeled DNA from *B. natto* were heated and annealed at 5 μg/ml. each in 1·9 × SSC. The lower tracing shows the three bands produced when the mixture is treated with the *E. coli* phosphodiesterase.

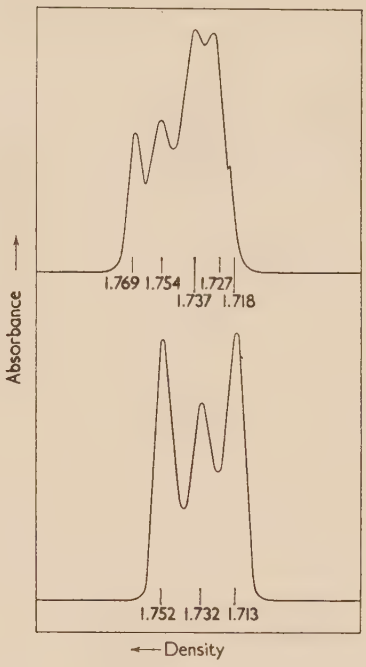


FIG. 12. Interspecies hybrid formation with *Sh. dysenteriae* DNA. Heavy-labeled *E. coli* DNA and unlabeled DNA from *Sh. dysenteriae* were heated and annealed at 5 μg/ml. each in 1·9 × SSC. When the mixture is treated with the phosphodiesterase and centrifuged in CsCl, only three bands appear.

the DNA of only one of the strains of *Salmonella* (*Salm. typhimurium*) has been tried but no hybrid was formed in this case either. We may then conclude that, in general, there seems to be no indication of sequence complementarity between the DNA of *Salmonella* and *E. coli* as measured by hybrid formation.

Thus in this case, where some homology was to be expected, none was found. It is likely that there is homology in some regions of the DNA molecules, but its failure to be displayed suggests that the homologous regions are dispersed or exist in only a few of the several hundred different molecules. Low concentrations of hybrid molecules would not be observed in the analytical ultracentrifuge, but could be isolated by using larger amounts of interacting DNA and working with the preparative swinging-bucket

rotor. In this way we plan to search very carefully for small amounts of hybrid that may be formed between ^{15}N-deuterated DNA from *E. coli* K12 and DNA from any one of the *Salmonella* strains.

Hybrid formation among other members of the family Enterobacteriaceae is now being investigated. Preliminary results with heavy *E. coli* B DNA and normal DNA from a strain of *E. freundii* (5610–52) or of *Erwinia carotovora* (ATCC 8061) show no

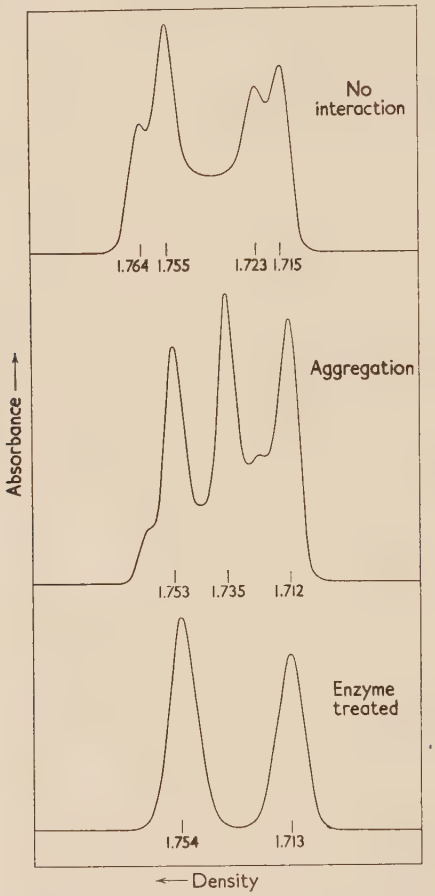


FIG. 13. Attempts at hybrid formation between *Salm. typhimurium* and *E. coli* DNA. The top tracing shows the results of heating and annealing DNA from *Salm. typhimurium* with ^{15}N-deuterated DNA from *E. coli*. Even at 10 μg/ml., which is double the usual concentration, no hybrids have been formed. In the central tracing an intermediate band appears as a result of heating at 20 μg/ml. each. This must be due to aggregates rather than double-stranded molecules since the band disappears after treatment with the phosphodiesterase (bottom tracing).

tendency toward hybrid formation. Unlabeled DNA isolated from another strain (17) of *E. freundii* did form hybrids with labeled *E. coli* B DNA. Since the classification of *E. freundii* is difficult, further studies with DNA of strains of this organism are necessary. Heavily-labeled DNA from *E. coli* K12, *Erwinia carotovora*, and *Salm. typhimurium* have also been prepared, allowing the study of other obvious possible relationships.

From this survey it can be concluded that equality of base composition and some genetic relation are necessary but not sufficient requirements for the formation of

hybrid DNA molecules composed of strands from different organisms. The present work is only indicative in nature but it does suggest that, with the further development of this technique, the degree of sequential homology between two different DNA molecules can be quantitatively explored.

9. Discussion

Until now we have been talking strictly in terms of the concept of strand separation and recombination, in accordance with the evidence presented by Doty *et al.* (1960). It has been seen that some important new facts can now be added to this evidence. By using heavy-isotope-labeled and normal DNA for heating and annealing experiments, it has been demonstrated that the units that unite in renaturation could not be the same as those that were united in the native DNA. It has also been demonstrated that the hybrids formed are not simply physically entangled aggregates. The strong species specificity of hybrid formation indicates that the links that hold the subunits of the hybrid molecules together are determined by the base sequence in each subunit. These links break whenever the hydrogen bonded structure of the molecule is broken. In fact, the rate of separation into subunits corresponds very closely to what has been predicted for the rate of unwinding of the strands of a double helix. It is possible that certain unknown links are broken whenever the Watson-Crick structure is disrupted and that the base sequence determines the specific type of linkage. The most natural explanation, however, seems to be that the strands of the Watson-Crick double helix do separate completely and come back together again through the formation of hydrogen bonds between base-pairs of complementary strands.

The hypothesis that genetic information resides in the sequence of bases in DNA has created a demand for techniques that will allow the study of the linear order of nucleotides along the DNA strand. An approach to this problem has been made in the nearest neighbor studies of Josse, Kaiser & Kornberg (1961) which offers a statistical evaluation of sequences displayed in DNA of various sources. By inference it can be assumed that genetic compatibility is also a measure of similarity in sequences between the DNA of the two parental strains. The present report has outlined a method by which it is possible to measure the extent of renaturation between homologous and heterologous DNA strands and has attempted to add a new dimension to the study of sequences of DNA of microbial origin.

The methodology of molecular hybrid formation *in vitro* has been outlined and its use in the study of similarities of DNA of various groups of micro-organisms has been illustrated. The close correlation between genetic compatibility, taxonomy, and hybrid formation has been mentioned. These applications are discussed in much greater detail elsewhere (Marmur & Schildkraut, 1961*b*; Marmur, Schildkraut & Doty, 1961). Organisms whose taxonomic classification is in doubt might readily be classified by first determining their base composition and studying their interaction by heating and annealing. Such experiments are now being extended to other members of the family Enterobacteriaceae (*Aerobacter*, *Klebsiella*, and *Serratia*) as well as the Pseudomonadaceae (*Pseudomonas Xanthomonas* and *Acetobacter*). Moreover, it might also be possible, by collecting homogeneous fractions of DNA of animal or plant origin, to investigate relationships in a similar manner. Aside from its taxonomic importance the technique offers a rational approach to the study of genetic compatibility where genetic exchanges have not yet been demonstrated.

It is a pleasure to express our gratitude to Dr. L. Grossman for his generous gift of the *E. coli* phosphodiesterase and for his valuable advice. We are especially indebted to Mr. William Torrey for his expert technical assistance with many aspects of this work, and to Mrs. F. Seelig for aid in preparation of the manuscript. We would also like to thank Drs. L. Wierzchowski and Donald M. Green, and Mr. Robert Rownd for their aid. The authors are very grateful to Professor C. E. Hall of the Massachusetts Institute of Technology for the electron micrograph. This investigation was supported by a grant (C-2170) from the National Cancer Institute, United States Public Health Service.

Note added in proof

It can be argued that the DNA species of intermediate buoyant density observed in the CsCl density gradient could be formed by non-specific, end-to-end aggregation of heavy isotope labeled and unlabeled DNA. It is evident, however, that biological hybrids could not be used to form DNA species possessing buoyant densities characteristic of either renatured fully labeled or fully unlabeled DNA unless strand separation and subsequent recombination of similarly labeled strands occurs during the heating and annealing procedure. Preliminary experiments by R. Rownd and D. Green using biological hybrids labeled with both ^{15}N and deuterium in only one strand and isolated from either *Escherichia coli* or *Bacillus subtilis* have shown that the heating and annealing procedure does result in renatured labeled and fully unlabeled, as well as hybrid, DNA molecules. The proportions are the same as observed when a mixture of labeled and unlabeled DNA is used as the starting material.

REFERENCES

Belozersky, A. N. (1957). In *The Origin of Life on the Earth.* Reports on the International Symposium, ed. by A. Oparin, p. 194. U.S.S.R.: Academy of Sciences.
Belozersky, A. N. & Spirin, A. S. (1960). In *The Nucleic Acids*, ed. by E. Chargaff & J. N. Davidson, Vol. III, p. 147. New York: Academic Press.
Cox, R., Marmur, J. & Doty, P. (1961). In preparation.
Davern, C. I. & Meselson, M. (1960). *J. Mol. Biol.* **2**, 153.
Doty, P., Marmur, J., Eigner, J. & Schildkraut, C. (1960). *Proc. Nat. Acad. Sci., Wash.* **46**, 461.
Dunn, D. B. & Smith, J. D. (1954). *Nature*, **174**, 305.
Eigner, J., Boedtker, H. & Michaels, G. (1961). *Biochim. biophys. Acta*, in the press.
Eigner, J. & Doty, P. (1961). In preparation.
Ephrati-Elizur, E. & Zamenhof, S. (1959). *Nature*, **184**, 472.
Ifft, J. B., Voet, D. H. & Vinograd, J. (1961). *J. Phys. Chem.*, in the press.
Josse, J., Kaiser, A. D. & Kornberg, A. (1961). *J. Biol. Chem.* **236**, 864.
Kuhn, W. (1957). *Experientia*, **13**, 301.
Lee, K. Y., Wahl, R. & Barbu, E. (1956). *Ann. Inst. Pasteur*, **91**, 212.
Lehman, I. R. (1960). *J. Biol. Chem.* **235**, 1479.
Lennox, E. S. (1955). *Virology*, **1**, 190.
Longuet-Higgins, H. C. & Zimm, B. H. (1960). *J. Mol. Biol.* **2**, 1.
Luria, S. E. & Burrous, J. W. (1957). *J. Bact.* **74**, 461.
Luria, S. E., Adams, J. N. & Ting, R. C. (1960). *Virology*, **12**, 348.
Marmur, J. (1961). *J. Mol. Biol.* **3**, 208.
Marmur, J. & Doty, P. (1959). *Nature*, **183**, 1427.
Marmur, J. & Doty, P. (1961). *J. Mol. Biol.* **3**, 585.
Marmur, J. & Lane, D. (1960). *Proc. Nat. Acad. Sci., Wash.* **46**, 453.
Marmur, J., Lane, D. & Doty, P. (1961). In preparation.
Marmur, J. & Schildkraut, C. L. (1961*a*). *Nature*, **189**, 636.

Marmur, J. & Schildkraut, C. L. (1961*b*). In *The Proceedings of the Vth Int. Congr. Biochem.*, London: Pergamon Press.
Marmur, J., Schildkraut, C. L. & Doty, P. (1961). In *The Molecular Basis of Neoplasia.* The Fifteenth Annual Symposium on Fundamental Cancer Research, ed. by S. Kit. Houston: Texas University Press.
Marmur, J., Seaman, E. & Levine, J. (1961). In preparation.
Marmur, J. & Ts'o, P. (1961). *Biochim. biophys. Acta*, in the press.
Meselson, M. & Stahl, F. W. (1958). *Proc. Nat. Acad. Sci., Wash.* **44**, 671.
Meselson, M., Stahl, F. W. & Vinograd, J. (1957). *Proc. Nat. Acad. Sci., Wash.* **43**, 581.
Rice, S. A. & Doty, P. (1957). *J. Amer. Chem. Soc.* **79**, 3937.
Rolfe, R. & Meselson, M. (1959). *Proc. Nat. Acad. Sci., Wash.* **45**, 1039.
Schildkraut, C. L., Marmur, J., Wierzchowski, L., Green, D. M. & Doty, P. (1961). In preparation.
Schildkraut, C. L., Wierzchowski, L. & Doty, P. (1961). In preparation.
Sinsheimer, R. L. (1959). *J. Mol. Biol.* **1**, 43.
Sueoka, N., Marmur, J. & Doty, P. (1959). *Nature*, **183**, 1429.
Szybalski, W., Opara-Kubinska, Z., Lorkiewicz, Z., Ephrati-Elizur, E. & Zamenhof, S. (1960). *Nature*, **188**, 743.
Watson, J. D. & Crick, F. H. C. (1953). *Nature*, **171**, 737.
Zamenhof, S. & Griboff, G. (1954). *Nature*, **174**, 306.
Zinder, N. D. (1960). *Science*, **131**, 813.

J. Mol. Biol. (1961) **3**, 618–633

X-Ray Diffraction Studies on the Large-Scale Molecular Structure of Insect Muscle

C. R. Worthington †

Medical Research Council Biophysics Research Unit, Wheatstone Physics Laboratory, King's College, Strand, London W.C.2, England

(*Received 13 April 1961*)

X-ray diffraction studies have been carried out on insect muscle, and it has been shown that the molecular filaments of both fibrillar and non-fibrillar muscle pack in a hexagonal lattice $L3$ with the actin filaments midway between the myosin filaments. A comparison with vertebrate striated muscle has verified that the molecular filaments of the latter muscle are packed according to lattice $L2$.

The observation of non-meridional low-angle X-ray reflections on a 410 Å layer line from insect muscle shows that a three-dimensional lattice is formed, which can arise only if there is lateral register between filaments. An attempt has been made to interpret this result in terms of structure; calculations on an analogue with cross-bridges between filaments arranged with a sixfold screw axis along the myosin filament axis are in agreement with the observed intensities.

1. Introduction

Recent work on the structure, histology and physiology of insect muscle and, in particular, insect flight muscle has been reviewed by Pringle (1957) and Boettiger (1960). The myofibrils of insect muscle have a cross-striation pattern similar to that of vertebrate muscle. Electron microscope studies have shown that one of the main structural differences between insect and vertebrate muscle is in the arrangement of molecular filaments in the myofibril.

In the electron microscope, transverse sections of vertebrate striated muscle show a double hexagonal array of different longitudinal filaments; each thin filament is surrounded by three thick filaments (Huxley, 1953*a*, 1957). The thin filaments are known to contain the protein actin, and the thick filaments contain the protein myosin (Huxley & Hanson, 1960). Therefore the filaments will be called the actin filaments and the myosin filaments, but this terminology does not exclude the possibility that the filaments contain other structural components. Studies on the same vertebrate muscle by X-ray diffraction (Huxley, 1953*b*) have given the inter-filament distances in intact muscle. Considerable shrinkage is known to occur in the fibril during specimen preparation for the electron microscope; X-ray studies are therefore required for more reliable information.

Similar electron microscope work on insect muscle (Huxley & Hanson, 1956) has shown that the double hexagonal array of filaments differs from vertebrate striated

† Present address: Department of Physics, University of Michigan, Ann Arbor, Michigan, U.S.A.

muscle in that each thin filament is midway between two thick filaments. The composition of the molecular filaments is not known for certain, but the appearance of the filaments in the electron microscope is rather similar to those of vertebrate striated muscle. The present X-ray results (to be described) have shown that the axial structure of the insect muscle filaments have the same periodicities as vertebrate striated muscle. Therefore the thin and thick filaments will be called the actin and the myosin filaments respectively.

Electron microscope studies on cross-striated muscle have demonstrated the presence of structures called cross-bridges between the actin and myosin filaments. The evidence for cross-bridges in vertebrate striated muscle has been reviewed recently by Huxley & Hanson (1960). The presence of cross-bridges in electron micrographs of insect muscle has been clearly shown by Hodge (1955) and also by Huxley & Hanson (1956).

In terms of current ideas on muscular contraction in cross-striated muscles, as reviewed by Huxley & Hanson (1960), it is supposed that certain linkages are made between the actin filaments and the myosin filaments. When these muscles shorten actively, they do so by a process involving the sliding of filaments past one another, and the linkages are in some way concerned with the resulting muscle tension. There is no direct evidence available on the nature of the linkages formed during contraction, but the presence of cross-bridges seen in electron micrographs corresponds in some way to these linkages.

The present X-ray studies on insect muscle have been undertaken in order to obtain information on the molecular structure of this muscle. The most detailed X-ray diffraction patterns were recorded from non-living muscle, although diffraction patterns of living relaxed muscle have also been recorded. So far X-ray diagrams have not been obtained from living muscle in the process of actively maintaining tension. It is realised that the present X-ray results (to be described) do not necessarily apply to the dynamic picture of muscular contraction, but that the X-ray studies, in common with the electron micrographs, refer to the molecular framework within which muscular contraction can take place.

2. Materials and Methods

The low-angle X-ray diffraction camera has been briefly described elsewhere (Worthington, 1959) and the specimen-to-film distance was either 5·6 cm or 6·7 cm. The camera could resolve up to 500 Å in the principal direction, which was just sufficient for the present work although a higher resolution would have been preferred. For the equatorial reflections a typical exposure was 12 hr.

The muscles used in this investigation were from the blowfly *Calliphora*, the water-beetle *Dytiscus*, the desert locust *Schistocera gregaria*, the lamellicorn beetle *Oryctes rhinoceros* and common varieties of bee and wasp from a garden in Surrey, England. In the case of the flight muscles the thorax was removed from the living animal and, usually, bisected longitudinally. The flight muscles were then fixed *in situ* or put into 50% glycerol or, if a fresh preparation was required, the muscles were removed from the thorax and put straight into the specimen chamber for investigation. In all cases, the flight muscles were only removed from the thorax immediately before X-ray examination.

The specimen holder was a Perspex container with two windows of thin polyester film. The dissected muscle was placed on filter paper which was moistened with insect Ringer for intact muscle or with fixative. The muscle remained in position by contact with the moist filter paper at each of its ends. The chamber was then made airtight by means of the cover plate and an O ring seal to prevent drying during exposure. The muscle specimen was aligned in the X-ray beam with the aid of a counter between the specimen holder and the X-ray film.

3. X-Ray Results

The low-angle X-ray diffraction pattern obtained from the double array of longitudinal filaments in insect muscle myofibrils has been indexed according to a three-dimensional hexagonal unit cell a, c. The two-dimensional hexagonal lattice vector $a = 560$ Å was obtained from intact muscle, from 50% glycerol-extracted muscle which was equilibrated with insect Ringer before examination and, also, from fixed or chemically treated muscle. The vector $c = 410$ Å lies along the axis of the molecular filaments. The ($hk\bar{i}l$) reflections are listed in Table 1 with their intensities on an arbitrary scale. So far only reflections on the equator ($l = 0$) and the first layer line ($l = 1$) have been recorded although traces of a second layer line ($l = 2$) have been seen. This detailed pattern has only been obtained from well orientated specimens and, in most cases, only the $10\bar{1}0$ and $20\bar{2}0$ reflections are observed.

Table 1

Equatorial reflections		1st layer line reflections	
$10\bar{1}0$	v.v.s.	$10\bar{1}1$	v.w.
$20\bar{2}0$	v.v.s.	$11\bar{2}1$	v.w.
$21\bar{3}0$	v.w.	$21\bar{3}1$	v.v.w.
$22\bar{4}0$	m.w.	$22\bar{4}1$	v.v.w.
$31\bar{4}0$	m.w.	$31\bar{4}1$	v.v.w.
$23\bar{5}0$	w.	$23\bar{5}1$	v v.w.

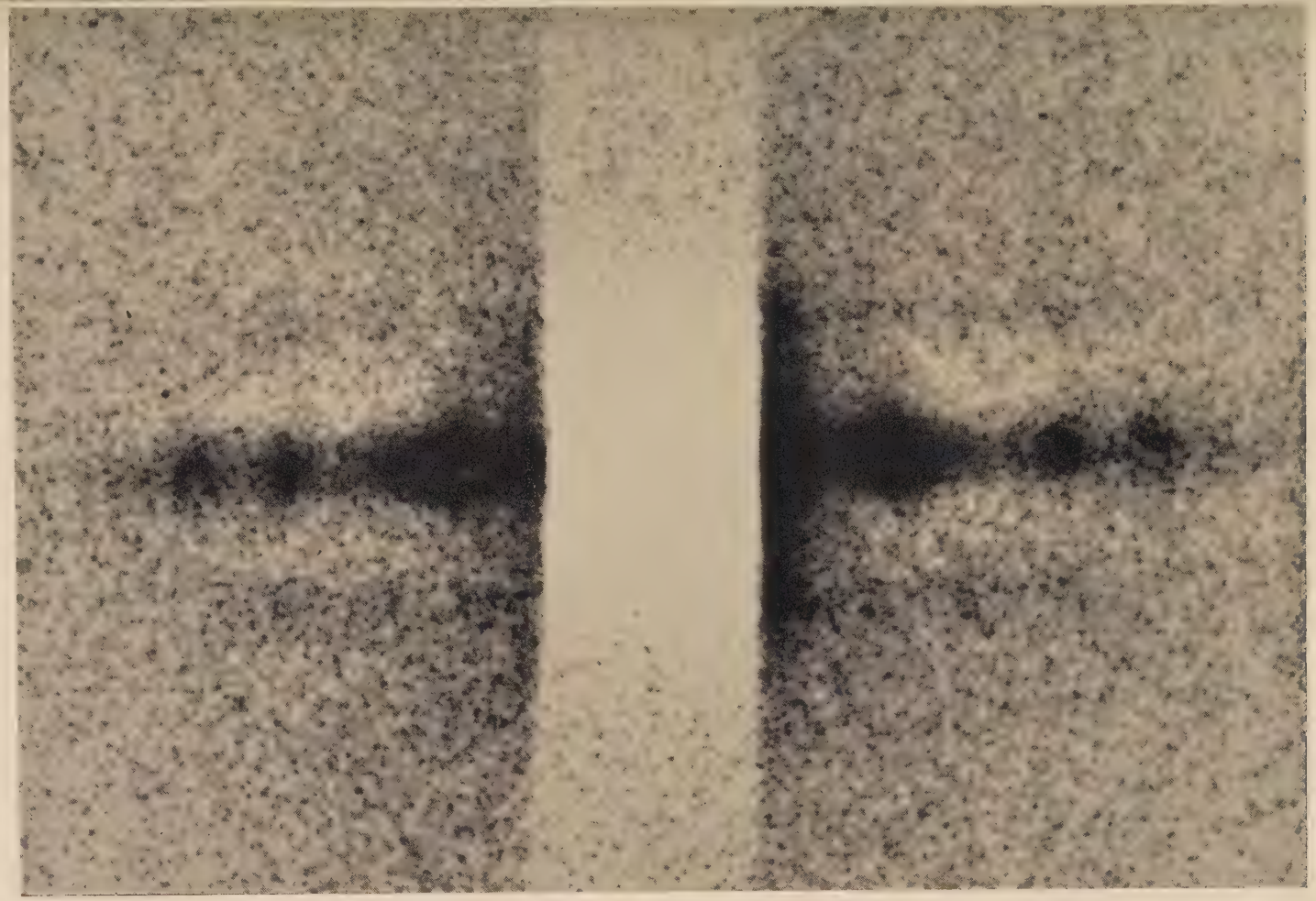
v., very; m., medium; s., strong; w., weak.

A photograph of this detailed pattern is shown in Plate I. The $10\bar{1}1$ reflection on the first layer line is close to the beam stop in the right half only. The other $hk\bar{i}1$ reflections are more evident in Plate II which shows one half of the X-ray pattern and its corresponding reciprocal lattice net in the opposite half. The X-ray pattern in Plate II represents the superposition of the four quadrants of Plate I and this representation brings out the weaker reflections. This method was suggested by Dr. M. H. F. Wilkins.

The weaker reflections were indexed by expressing the spacings as fractions of the $10\bar{1}0$ spacing. It can be seen from the reciprocal lattice net that the $22\bar{4}0$ and $31\bar{4}0$ reflections are close together, with fractions of 0·289 and 0·277 respectively, and these reflections have not been resolved. An average fraction of 0·282 was obtained from numerous patterns. This fraction has been observed as 0·288 and 0·278 in different patterns, which suggests that the $22\bar{4}0$ and $31\bar{4}0$ reflections respectively contribute strongly to the observed patterns. Slight changes in the lattice vector a will no doubt introduce changes in the intensity of the reflections.

4. Fibrillar and Non-fibrillar Muscle

Insect muscle has been classified histologically as either fibrillar or non-fibrillar, according to the size of the myofibrils and the way the myofibrils are arranged in the



[*To face page* 620

PLATE I. The low-angle X-ray diffraction pattern from the flight muscles of a blowfly. Specimen-to-film distance 5·6 cm.

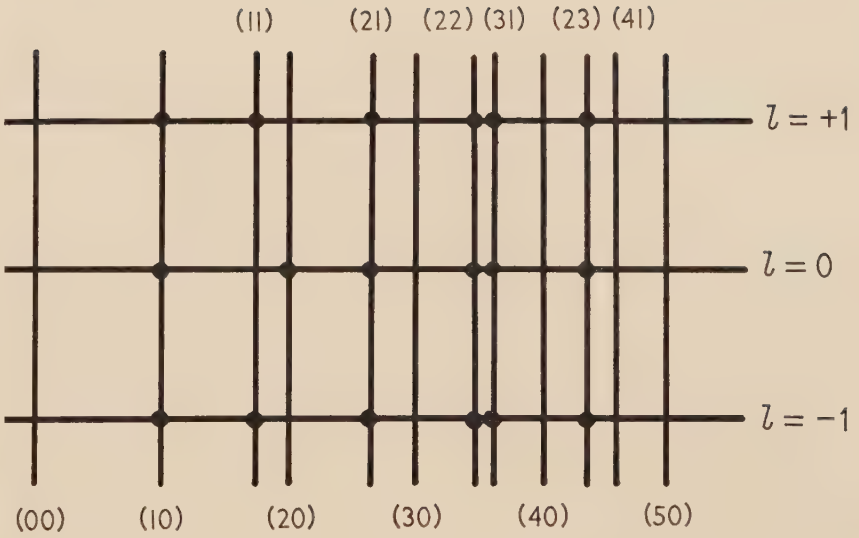


Plate II. In the left half the X-ray pattern is that obtained by superimposing the four quadrants of Plate I. The right half shows the corresponding hexagonal reciprocal lattice net. The grid consists of horizontal lines of constant l and vertical lines of constant hk. The intersection points are possible reflections; the observed reflections are marked as solid circles. The (00) line is correctly placed with respect to the X-ray pattern.

muscle fibre (Pringle, 1957; Boettiger, 1960). Fibrillar muscle has a very high frequency of operation which exceeds the excitation frequency, whereas non-fibrillar muscle has a 1:1 relation between its operation and the nerve impulse.

The fibrillar muscles examined were the flight muscles of the blowfly, waterbeetle, lamellicorn beetle, bee and wasp. Usually the dorsal longitudinal muscle was chosen, but on occasions the dorsoventral muscle was also chosen. In the case of the lamellicorn beetle, the basalar muscle was examined because the mechanical properties of this particular muscle have been recently studied (Machin & Pringle, 1959). The non-fibrillar muscles include the flight muscles and the jumping leg muscle of the desert locust ("close-packed muscle") and the leg muscles of the waterbeetle ("tubular muscle").

All the muscle specimens examined have given $10\bar{1}0$ and the $20\bar{2}0$ X-ray reflections, with the same lattice vector $a = 560$ Å. Therefore the molecular filaments are arranged in the same way in the myofibrils of both fibrillar and non-fibrillar insect muscle, and have the same inter-filament separations.

5. Living and Dead Muscle†

Low-angle X-ray studies of Huxley (1953*b*) on living vertebrate muscle in a resting condition, and the wide-angle X-ray studies of Astbury (1947) on living relaxed muscle of various kinds, have shown no difference in X-ray spacings between living and dead muscle. The question arises whether living and dead muscles have the same lattice spacings. A partial affirmative answer can be given because freshly prepared insect muscle always gives the same X-ray reflections as those from dead muscles, even with exposure of a few hours. It is known from physiological studies of insect flight muscle that isolated muscles can remain alive for up to 8 hr (Weis-Fogh, 1956; Machin & Pringle, 1959).

An even more definite answer to this question is given by the observation of the low-angle equatorial X-ray diffraction pattern from the intact flight muscles of a live blowfly. The blowfly was held with plasticene in the specimen holder and the X-ray beam was passed through the thorax in the vicinity of the dorsal longitudinal muscle. An exposure of 5 hr was sufficient to record the X-ray reflections even allowing for the extra absorption of the other material in the thorax. (Although the low-angle camera produces an intense X-ray beam, the total number of X-ray quanta in the beam is quite small and, in practice, the blowfly will remain alive for at least 1 day.) The X-ray spacings are the same as those of dead muscles, so that both living and dead muscles have their filaments arranged in the same two-dimensional hexagonal lattice $a = 560$ Å.

In most studies on the ultrastructure of muscle it has been assumed that living and dead muscles have essentially the same filament packing, as is now confirmed by the X-ray observations. Fischer (1947) has shown that living and dead muscles have identical birefringences when the fixative is correctly chosen. Also the fact that form birefringence has been distinguished from total birefringence with reproducible results, on the basis of the Weiner theory of protein rodlets, is evidence that the molecular filaments maintain a constant filament separation.

† Living muscle, as defined here, can shorten actively on electrical stimulation whereas dead muscle cannot. Dead muscles, as defined here, are those in rigor, glycerol-extracted muscles and fixed muscles.

6. Comparison with Vertebrate Muscle

A brief report on the transverse packing of filaments in both vertebrate and insect muscle has been published already (Worthington, 1960) together with observations on swelling and drying.

The X-ray studies of vertebrate muscle by Huxley (1953*b*) recorded the $10\bar{1}0$ and $11\bar{2}0$ reflections with the lattice vector $a = 440$ Å. Insect muscle has a lattice vector $a = 560$ Å (Worthington, 1960). The two-dimensional hexagonal lattice of vertebrate muscle was denoted *L*2 as there are 2 actin filaments per cell unit, while the insect muscle lattice was denoted *L*3 since there are 3 actin filaments per unit cell. In *L*2 the actin filaments are at $\frac{2}{3}$, $\frac{1}{3}$ and $\frac{1}{3}$, $\frac{2}{3}$ while those in *L*3 are at $\frac{1}{2}$, 0; 0, $\frac{1}{2}$ and $\frac{1}{2}$, $\frac{1}{2}$. These arrangements are illustrated in Fig. 1.

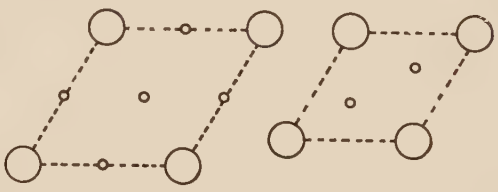


Fig. 1. Diagrams of the two-dimensional unit cells of lattices *L*2 (right) and *L*3 (left). The larger circles represent myosin (120 Å diameter) and the smaller circles represent actin (50 Å diameter).

Electron microscope studies determine directly the number and position of the actin and myosin filaments in the lattices *L*2 and *L*3. The simplest indexing of the equatorial X-ray reflections defines the primitive lattice vector *a*. The observation of the strong $hk\bar{i}0$ reflections can establish the position of the actin filaments without reference to the electron micrographs; however, it is more convenient to assume the positions suggested by electron microscopy and to use these to account for the observed intensities.

Approximate calculations can be made by assuming that the actin and myosin filaments are long cylinders of uniform density with diameters of 50 Å and 120 Å respectively. Effective temperature factors are expected to be large, but these have been omitted from the calculations, and each filament is assumed to lie along the *c* direction with no displacement perpendicular to its length. The equatorial intensities refer to muscles at body length; vertebrate muscle has the lattice vector $a = 440$ Å and insect muscle has $a = 560$ Å. The intensities I $(hk\bar{i}0)$ have been computed on an arbitrary scale for lattices *L*2 and *L*3 and are given in Table 2.

Table 2

L 2		*L* 3	
I $(10\bar{1}0)$	10, observed	I $(10\bar{1}0)$	10, observed
I $(11\bar{2}0)$	10, observed	I $(11\bar{2}0)$	3, not observed
I $(20\bar{2}0)$	1, not observed	I $(20\bar{2}0)$	10, observed

I $(hk\bar{i}0)$, calculated intensity on an arbitrary scale.

It can be seen that the calculated intensities of the observed reflections are large, from which it follows that the X-ray data are consistent with the electron micrographs.

The separation of the actin and myosin filaments in the two lattices may be obtained as follows. The centre-to-centre distance is 254 Å for *L*2 and 280 Å for *L*3 and, if we assume diameters of 120 Å for myosin and 50 Å for actin, the surface-to-surface distances are 169 Å for *L*2 and 195 Å for *L*3. The difference between these figures might in fact be less, for there is evidence from electron microscopy that the myosin diameter may be slightly larger in insect muscle.

7. Swelling and Drying

One of the difficulties in the study of water uptake and removal in muscle is the correct assignment of water and other substances to extra- and intracellular compartments (Harris, 1960). Both vertebrate and insect muscle myofibrils contain about 90% fluid by volume. It has not so far been possible by controlled drying at different humidities to produce a lattice vector a less than 560 Å in insect muscle; drying destroys the regular arrangement of filaments and no discrete equatorial diffraction is observed. It is of interest that the regular order can be regained on rehydration, the lattice vector being unchanged.

It is known that muscle, as well as other tissues, can swell in most fluids and that the amount of fluid uptake depends on various factors such as the charge on the filaments (in the case of muscle) and the ionic media (Harris, 1960). Various treatments, such as distilled water, 15% acetone and 1% phosphotungstic acid, cause the insect muscle to swell considerably. The equatorial reflections $10\bar{1}0$ and $20\bar{2}0$ are still observed but the lattice vector a becomes greater. Thus, when muscle swells under the above conditions the arrangement of filaments in the myofibril is maintained with an increase in the actin-to-myosin separation.

The lattice vector a has been increased by at least 50% for *L*3; this involves an increase in the myofibril cross-section of 125% and an increase in the inter-filament distance (actin-to-myosin) of about 100 Å. It is assumed that the filaments themselves do not increase in size during swelling. This behaviour of swelling in muscle can be compared to that of various gels where long-range repulsive forces which are electrostatic in origin maintain a regular structure with an average inter-molecular distance as large as 1000 Å. Probably these long-range forces play an important role in muscular contraction, and it is suggested that the filament lattice of both living relaxed muscle and swollen muscle is maintained almost solely by these forces.

8. The Axial Period of Striated Muscle

Electron microscope studies (Draper & Hodge, 1949) have shown a regular striation transverse to the fibre axis with a period of about 400 Å in both the A and I bands of striated muscle. Improved methods using thin sections of muscle have demonstrated that there are cross-bridges between the actin and myosin filaments. It follows that both the structure of the myosin and actin filaments as well as the cross-bridges between them could make contributions towards the 400 Å axial period of muscle. The actin and myosin filaments show no structure in the electron microscope but the arrangement of the cross-bridges between filaments in vertebrate muscle has been studied in some detail by Huxley (1957).

Huxley (1957), in a high resolution electron microscopy study on vertebrate muscle (rabbit psoas muscle), has described the cross-bridges in terms of lateral projections which originate from the myosin filaments alone. This assignment comes directly from Huxley's observations that, in stretched muscle, projections can be seen on the myosin filaments in that part of the A band from which the actin filaments have been withdrawn: the actin filaments in the I band have no projections. There are six projections from each myosin filament every 400 Å, and the projections are arranged on a sixfold screw axis, so that the actin filaments which they connect would have a threefold screw axis every 400 Å.

In electron micrographs of both insect muscle (Huxley & Hanson, 1956) and vertebrate muscle (Huxley, 1957) cross-bridges appear to be in contact with both the actin and myosin filaments and lie roughly at right angles to the filament axis. Cross-bridges between filaments could be some macromolecular component attached to or forming part of either the myosin filament, or the actin filament, or both. Such macromolecular components would give the appearance of cross-bridges between filaments in the electron microscope. Specimen preparation for electron microscopy usually involves fixation with osmic tetroxide, dehydration with alcohol, further staining with phosphotungstic acid, and embedding in methacrylate or in Araldite before sectioning. The muscle specimens invariably shrink during this treatment: this is shown by the estimated value of the lattice vector a of vertebrate muscle in electron micrographs which is about 300 to 350 Å (Huxley, 1957), i.e. 100 to 150 Å smaller than the value observed in live muscle. Therefore, the electron microscope observation that the lateral projections from the filaments are in contact, and hence form cross-bridges, is not necessarily valid for live muscle. The X-ray observations on swollen muscle are perhaps relevant here, for the actin-to-myosin distance has been increased by up to 100 Å without any breakdown in the filament lattice. The concept of a separation between the lateral projections of the filaments will be discussed in more detail later.

The observation with the electron microscope of an axial period of about 400 Å in rather thick sections of striated muscle suggests that there is a considerable degree of lateral register between filaments. In insect muscle a three-dimensional lattice a, c has been observed by means of X-ray diffraction. A three-dimensional lattice is only possible when projections are in register from one filament to the equivalent filament in the next unit cell in a direction perpendicular to the filament axis.

The existence of the three-dimensional lattice of insect muscle would seem to be firmly established because the diffraction pattern can be indexed according to a three-dimensional unit cell a, c. The alternative concept that either the actin filaments or the myosin filaments of insect muscle have projections or ridges with a spacing of 410 Å which diffract as independent elements and give rise to the series of reflections on the 410 Å layer line can be discounted as the reflections are dependent on the lattice vector a. By swelling insect muscle the value of a increases and the reflections on the 410 Å layer line move correspondingly.

The three-dimensional lattice of insect muscle could be formed by a lateral register of either the myosin filaments, or the actin filaments, or both. It is known from previous X-ray studies that the myosin filaments in vertebrate striated muscle have an axial period of 435 Å (Worthington, 1959) and that the actin filaments have an axial period of 410 Å (Selby & Bear, 1956; Worthington, 1959). As mentioned previously there is evidence that the actin and myosin filaments of insect muscle have the same axial periodicities as vertebrate striated muscle. It is thought likely that

the projections on the filaments would be spaced at intervals equal to the axial structure of the filaments. There are, therefore, two possible lattices: that formed by lateral register of myosin filaments and with vector $c = 435$ Å, and the lattice formed by actin filaments in register and with $c = 410$ Å.

To what extent the results of electron microscopy apply to this observation of a three-dimensional lattice by X-ray diffraction is a matter of debate. X-ray results refer to intact (live) muscle whereas the muscles in the electron microscope have undergone considerable treatment during the specimen preparation previously described. Nevertheless, one would expect the observed diffraction patterns to relate to structural features which are apparent in the electron microscope. The electron microscope observations strongly support the existence of projections on the myosin filaments but gives only weak support, if any, to the existence of projections on the actin filaments. Hence these considerations favour a three-dimensional lattice formed by the lateral register of the myosin filaments.

9. X-ray Analysis

The low-angle X-ray diffraction pattern from insect muscle is not as precise or as detailed as many diffraction patterns, for there are relatively few reflections where the intensities are known even qualitatively. Nevertheless, it is clear that the observed low-angle X-ray reflections arise from a three-dimensional unit cell $a = 560$ Å, $c = 410$ Å. It remains to be seen whether a model can be proposed to fit the X-ray data in Table 1 as well as other known results from electron microscopy and X-ray diffraction. A proposed model for muscle which incorporates only large-scale structural features for purposes of interpreting the low-angle X-ray diffraction pattern should perhaps be called an analogue and this term will be used here. An analogue will have to take account of both the myosin filaments and the actin filaments as well as the cross-bridges between them. Each of these three factors will be considered in turn in order to determine their relative contribution to the observed $hk\bar{\imath}l$ reflections. For the present purposes the observed intensities of the X-ray reflections given in Table 1 can be compared with the calculated values of $|F|^2$ where F is the structure factor of the unit cell contents. $|F|^2$ is equal to $(\Sigma_j f_j A_j)^2 + (\Sigma_j f_j B_j)^2$, f is the scattering factor, A and B are the usual geometric structure factors as defined in the *International Tables for X-ray Crystallography*, Volume 1, and the summation over j includes contributions from both the myosin and the actin filaments and from cross-bridges between them.

The analogue used in the calculations is illustrated in Fig. 2. The fibre axis and axes of the molecular filaments lie along the z axis, the myosin and actin filaments have diameters of 120 Å and 40 Å, respectively, with electron density ρ a function of z only, i.e. $\rho = \rho(z)$, and the electron density function repeats at intervals of 435 Å and 410 Å respectively. A diameter of 40 Å has been chosen for the actin filaments rather than 50 Å solely to simplify the calculations. The cross-bridges in electron micrographs of insect muscle, e.g. Huxley & Hanson (1956), appear to have about the same cross-sectional area as the actin filaments; hence it would be natural to choose a cylinder of length 200 Å and diameter 40 Å to represent each cross-bridge. However, for ease of calculation the cross-bridge has been represented by five spheres of 40 Å diameter which are in contact with each other and with the myosin and actin filaments. The calculations are closely similar when spheres of smaller diameter and in larger numbers are chosen to represent the cross-bridges.

The analogue construction depends on whether space group $P6$, $P6_1$ or $P6_5$ is chosen. Calculations involving space group $P6_1$ and $P6_5$ are identical so that only $P6_1$ will be mentioned hereafter. In space group $P6$ the z coordinates of the six cross-bridges are all $z = 0$; the six cross-bridges lie in the same plane at right angles to the filament axis. In space group $P6_1$ the z coordinates of the six cross-bridges are at intervals of $\frac{1}{6}$ of the unit cell length (410 Å); the six cross-bridges are arranged helically with a sixfold screw axis along the c direction which is coincident with the myosin filament axis.

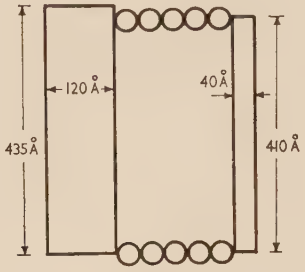


FIG. 2. Diagram of the muscle analogue on which the X-ray calculations are based, showing part of the $a0c$ plane. The a direction is that of the two-dimensional lattice and the c direction is vertical and lies along the fibre axis.

The X-ray crystallographic calculations are summarized in the Appendix. In the consideration of the $hk\bar{i}0$ reflections the calculations are the same for both space groups $P6$ and $P6_1$. The actin and myosin filaments contribute to all reflections and the calculations for the first three reflections are given in Table 2. The cross-bridges also contribute to each $hk\bar{i}0$ reflection. However, the contributions oscillate from positive to negative values and are relatively small, especially as the scattering factor of a sphere of diameter 40 Å is much less than the scattering factor of the actin and myosin cylinders of length 410 Å. It is of interest that the contributions from the cross-bridges would increase the values of the $10\bar{1}0$ and $20\bar{2}0$ reflections and decrease the value of the $11\bar{2}0$ reflection given in Table 2. No conclusions can be drawn concerning the other $hk\bar{i}0$ reflections as knowledge of the actual scattering factors would be required.

The consideration of the $hk\bar{i}1$ reflections require the space groups $P6$ and $P6_1$ to be distinguished.

The myosin filaments cannot make any contribution as we are summing a 435 Å repeat structure at intervals of 410 Å with the result that the contributions will be zero on the average. In space group $P6$ the actin may have a geometric contribution and it is expected that reflections with h, k even, i.e. $00\bar{0}1$, $20\bar{2}1$, $22\bar{4}1$ and $40\bar{4}1$, would be strongest. However, in Table 1 we only observe the $22\bar{4}1$, $31\bar{4}1$ reflection. The cross-bridges give a very strong $00\bar{0}1$ contribution and this reflection is not observed. It is concluded that an analogue based on space group $P6$ is therefore almost certainly eliminated.

In space group $P6_1$ only the cross-bridges can contribute to the $hk\bar{i}1$ reflections and large $(A^2 + B^2)$ values are obtained for the $10\bar{1}1$, $11\bar{2}1$, $21\bar{3}1$ and $31\bar{4}1$ reflections; possibly three out of these four reflections are observed. It is concluded that space group $P6_1$ is more likely.

It follows that the analogue (illustrated in Fig. 2) is constructed in accordance with space group $P6_1$ and has a hexagonal unit cell $a = 560$ Å, $c = 410$ Å. There are two strong equatorial reflections principally deriving from the actin and myosin filaments, and a series of reflections ($l = 1$) which arise from the cross-bridges between the filaments. Agreement between calculated and observed intensities are as good as can be expected considering the nature of the problem.

A test of the analogue can be made to determine whether X-ray scattering from the cross-bridges can account for the intensity of the layer line reflections compared with the equatorial reflections. The $10\bar{1}0$ and the $10\bar{1}1$ reflections were chosen and the observed intensity ratio $I(10\bar{1}0)$: $I(10\bar{1}1) = 200 : 1$ was estimated using the most detailed pattern so far obtained; most patterns have a much larger ratio of observed intensities. In order to compare the observed intensity ratio of 200 : 1 with that calculated from the analogue the $|F|^2$ values of the unit cell a, c were computed assuming each component had the same electron density. The ratio

$$|F(10\bar{1}0)|^2 : |F(10\bar{1}1)|^2 = 60 : 1$$

was obtained using all components of the analogue for the $10\bar{1}0$ reflection and only the cross-bridges for the $10\bar{1}1$ reflection. The calculated intensity ratio $I(10\bar{1}0) : I(10\bar{1}1)$ will, in general, be greater than the observed $|F|^2$ ratio because the intensities of the first layer line reflections depend on the percentage of the cross-bridges in lateral register, whereas the intensities of the equatorial reflections are independent of whether the filaments are in lateral register or not. It is concluded that the observed intensity ratio of the equatorial and layer line reflections is in agreement with that calculated from the analogue, and therefore the X-ray scattering from the cross-bridges is sufficient to account for the observed intensity of the layer line reflections.

10. Correlation with Other X-Ray Diffraction Results

Insect muscle gives the same wide-angle X-ray diffraction pattern as vertebrate striated muscle, molluscan smooth muscle and myosin (Astbury, 1947). This is the well known α-pattern. The thick filaments of muscle, the myosin filaments of vertebrate striated and insect muscle, and the paramyosin filaments of molluscan smooth muscle, all give rise to the wide-angle pattern. The actin filaments do not contribute to the wide-angle pattern but give a series of moderate-angle reflections. In a previous X-ray study (Worthington, 1959) it was demonstrated that the myosin filaments of vertebrate striated muscle give a series of low-angle X-ray reflections. Two X-ray spacings were carefully measured and were shown to correspond to a series of low-angle reflections from living relaxed frog sartorius muscle which had been attributed to the actin filaments (Huxley, 1953*b*). These two X-ray spacings were the 3rd and 6th order of an axial period of 435 Å which was attributed to the myosin filaments. Later work has shown additional orders in the case of rabbit psoas muscle (Worthington, unpublished results) and a well defined pattern has been obtained by Elliott (1960) from living relaxed frog sartorius muscle showing the 2nd, 3rd, 4th and 6th orders of diffraction. These patterns were obtained with two-dimensional resolution and showed that all orders of diffraction were meridional and consistent with an axial repeat distance of 435 Å.

In the case of insect muscle only one reflection has been recorded from the myosin filaments. This reflection, at 145 Å, has only been observed after the muscle was

"stained" with 1% phosphotungstic acid. Additional orders of diffraction are necessary if the axial period is to be established unambiguously; however, in correspondence with vertebrate striated muscle, the myosin filaments of insect muscle appear to have the same axial period of 435 Å.

The moderate-angle actin reflections from various kinds of air-dried molluscan smooth muscles and from vertebrate striated muscles are known to be identical (Worthington, 1959). These reflections have been correlated with a helical structure of 15 residues in 7 turns which has a repeat distance of 410 Å (Selby & Bear, 1956; Worthington, 1959). Air-dried insect muscle gives the same actin pattern as the above muscles and therefore the actin filaments of insect muscle have an axial period of 410 Å.

The axial periods of the myosin and the actin filaments of cross-striated muscle are known to be 435 Å and 410 Å, respectively, with an accuracy of 1%. In the case of vertebrate striated muscle a further meridional reflection at 410 Å was identified and measured to an accuracy of 2%. This reflection was not part of either the actin or the myosin patterns. Also it was shown that the spacing of this reflection could be reduced by chemical treatment (Worthington, 1959).

The lattice vector a of insect muscle has already been compared with that of vertebrate muscle, and the actin-to-myosin centre-to-centre distance has been shown to be about the same in both lattices. The accuracy of the measurements is 2%, i.e. $a = 560 \pm 10$ Å.

The determination of the magnitude of the lattice vector c was difficult, for there was no internal standard; the lattice vector a could not be used because it varied. Also the reflections were not particularly well resolved and the distance to be measured on the X-ray film was small (about $\frac{1}{2}$ mm). The accuracy of an individual measurement was about 3% for the best patterns. The average value obtained for c from ten low-angle patterns of insect muscle showing the meridional layer line reflections was 403 ± 3 Å. Some of the specimens included in this average were fixed muscles which had a slightly smaller value of c. The meridional reflection of vertebrate muscle at 410 Å behaves similarly and the decrease in spacing is larger than that observed with insect muscle. Allowing for this factor the lattice vector c of insect muscle would approach the value 410 Å, i.e. $c = 410 \pm 10$ Å.

Because the lattice vector c has the same magnitude as the repeat period of the actin filaments, it is of interest to describe in some detail the structure of the actin filaments. In the electron microscope actin filaments appear to have diameters of about 50 Å. X-ray determinations of the actin diameter in muscle give values of approximately 50 Å (Selby & Bear, 1956) or approximately 60 Å (Worthington, unpublished results) when it is assumed that the moderate-angle X-ray reflections of actin come from a helical structure. The X-ray measurement of a 50 to 60 Å diameter for actin in muscle suggests that this value applies to the narrow rod structure assumed in the analogue and gives no information concerning possible projections from the actin filaments which could contribute to the cross-bridges. X-ray patterns of reconstituted actin fibres have been obtained by Cohen & Hanson (1956). These fibres very likely consisted of actin filaments packed together in a crystalline lattice. The filament separation was about 55 Å. This suggests that the actin filaments had a narrow rod structure and possible projections were missing from the reconstituted filaments.

11. Discussion

An analogue for insect muscle has been put forward to explain the low-angle X-ray diffraction pattern. This analogue includes the large-scale features of molecular structure which have been obtained directly from X-ray diffraction studies. The cross-bridges are represented by a row of spheres. The bridges are in contact with both the actin and the myosin filaments, in agreement with the results of electron microscopy. Calculations involving variation in the number of the spheres between the filaments have been useful in that some extremes can be eliminated. The results of such calculations suggest that the layer line reflections are unlikely to arise from some small structure on either the actin filament or the myosin filament. A structure such as a small sphere on the actin filament would have a co-ordinate $x = \frac{1}{2} \pm \Delta x$ where Δx is small. The contributions from these structures are very small for small hk and this is contrary to what is observed. The contributions from similar small structures on the myosin filaments are of reasonable size but have an intensity distribution which differs markedly from that shown in Table 1. Apart from these two extreme examples, it is not possible with the present X-ray data to improve on the analogue, or to specify the projections in greater detail. It has also been shown that the X-ray scattering from the cross-bridges represented by the row of spheres is sufficient to account for the observed intensities of the layer line reflections. This would indicate that the cross-bridge structure in the analogue corresponds to the cross-bridge structure seen in electron micrographs.

As previously mentioned, cross-bridges could arise from projections on the filaments, and it is thought likely that there is a distance separating the lateral projections in live muscle. In order to explain the contractile force between the actin and myosin filaments Spencer & Worthington (1960) have suggested a mechanism according to which there are electrostatic forces between the filaments. The forces are exerted between sites spaced at intervals of 410 Å and 435 Å on the actin and myosin filaments respectively. The muscle contracts when a system of charge transfers passes down the actin filament from the Z-line. The projections effectively reduce the distance between filaments so that the resultant force will be greater. The effective gap between filaments is thought to be much less than the surface-to-surface distance of about 200 Å. This can be represented in Fig. 2 by the removal of one or more of the spheres and adjusting the remaining spheres on the myosin filaments so that they have a repeat distance of 435 Å. In this case, only the spheres on the actin can contribute to the 410 Å layer line reflections, for the reasons given before.

The cross-bridges in the analogue have been arranged according to space group $P6_1$ and they can be described in terms of either a sixfold screw axis along the c axis or a twofold screw axis parallel to c and passing through $(\frac{1}{2}, 0)$. This description of the cross-bridges is independent of the composition of the cross-bridges and requires the myosin filament to have a sixfold screw axis and the actin filament to have a twofold screw axis. In accordance with the electron microscope observations of Huxley (1957) on rabbit psoas muscle, the projections on the myosin filaments are arranged with a sixfold screw axis and the actin filaments which they connect have a threefold screw axis. Therefore, the myosin filaments of both insect muscle and vertebrate striated muscle have a sixfold screw axis, and the actin filaments in insect muscle have a twofold screw axis whereas those in vertebrate striated muscle have a threefold screw axis. It can be seen in Fig. 1 that the myosin filaments in lattices $L2$ and $L3$

have a sixfold rotation axis along c, whereas the actin filaments in $L2$ have a twofold rotation axis parallel to c and passing through $(\frac{1}{2}, 0)$ and those in $L3$ have a threefold rotation axis parallel to c and passing through $(\frac{2}{3}, \frac{1}{3})$. The two-dimensional symmetry involving rotation axes in $L2$ and $L3$ depends on the way the actin and myosin filaments are arranged in the hexagonal lattice whereas the three-dimensional symmetry involving screw axes in $L2$ and $L3$ arises from the helical arrangement of cross-bridges.

The myosin filaments of vertebrate striated muscle give a series of meridional reflections which correspond to an axial structure repeated along the axis of the myosin filament at 435 Å. By itself this gives no information about the symmetry of the myosin filament. The actin filaments have been described in terms of a helical structure of 15 residues and 7 turns. This structure has a threefold screw axis, but does not possess a twofold screw axis. If the same actin filament was present in both lattices $L2$ and $L3$, then the filament would be required to have a sixfold screw axis. Even if one assumes that a particular residue of the actin structure either connects, or attracts, a projection from the myosin filament or, alternatively, if a projection is in fact attached to, or part of, that residue, then the symmetry of the actin filament presents a difficulty which has not been resolved. However, this topic does not directly relate to the important question whether the cross-bridge structure arises from projections on either the actin filaments, or the myosin filaments, or both, because the symmetry of the actin and myosin filaments in lattice $L2$ and $L3$ is independent of the composition of the cross-bridge structure.

The X-ray measurement of the lattice vector c has given the repeat distance of the helically arranged cross-bridge structure which contributes to the layer line reflections. As can be judged from Plates I and II, the reflections ($l = 1$) are not well defined, and therefore the measurement of the layer line spacing is not very precise. The lattice vector c is given above as 410 Å with an accuracy of $2\frac{1}{2}\%$ whereas the axial periods of 435 Å and 410 Å are known to 1%. The magnitude of this repeat distance is the same as the axial period of the actin filaments. This suggests that the cross-bridge structure which contributes to the layer line reflections comes from projections associated with the actin filaments. Therefore, on the basis of the known axial periods of actin and myosin filaments, the assignment of projections in insect muscle can only be made to the actin filaments.

It has been shown by X-ray diffraction that the myosin and actin filaments of insect muscle have the same periodicities as vertebrate striated muscle. Therefore, it can be argued that because, according to the electron microscope study, vertebrate muscle has projections originating from the myosin filaments, then insect muscle should also have projections on the myosin filaments. There is no X-ray evidence for projections spaced at intervals of 435 Å. This could be because myosin in insect muscle diffracts weakly or because the myosin filaments are not in lateral register. The helically arranged cross-bridge repeat distance has been recorded at 410 Å. If this structure originated from the myosin filaments, then the projections at 410 Å would not be an obvious part of the axial structure at 435 Å. Alternatively, projections from the myosin filaments every 435 Å might be attracted towards the sites of the actin filaments, and this could give a pseudo 410 Å repeat distance. It is clear from the above that the question whether the projections of insect muscle come from the actin filaments, from the myosin filaments, or from both, cannot be settled definitely by the present X-ray data alone.

The X-ray analysis has shown that the layer line reflections of insect muscle could arise from projections arranged helically. The low-angle X-ray diffraction pattern had a number of non-meridional reflections; the meridional reflection 00$\bar{1}$1 was missing. X-ray diffraction theory requires a helical arrangement of projections to give non-meridional reflections on the first layer line, and a meridional reflection can only arise from a non-helical arrangement of projections. In the case of vertebrate muscle, apart from the diffraction from the filaments themselves, the only reflection observed was the 410 Å meridional reflection (Worthington, 1959). This reflection was clearly shown to be meridional, because the low-angle X-ray diffraction patterns had two-dimensional resolution. The spacing was the same as both the axial period of the actin filaments and the layer line repeat distance of insect muscle. Therefore it is reasonable to interpret this reflection as arising from a structural component associated with the actin filaments. This component may or may not be identical to the projections of insect muscle. The 410 Å meridional reflection of vertebrate muscle cannot arise from projections which are helically arranged. If the diffraction comes from projections on the actin filaments, then the observation of this diffraction contradicts the helical arrangement of cross-bridges deduced from electron microscopy. Alternatively, if the cross-bridges are helically arranged, then the diffraction arises from an axial component along the actin filaments, or less likely, from a disordered array of cross-bridges. The additional observation that the magnitude of this X-ray spacing can be decreased by chemical treatment without any change in the myosin reflections, but with the disappearance of the actin pattern, would suggest that this reflection is directly related to the actin filaments.

The observation of two non-net reflections with an axial period of about 400 Å in molluscan smooth muscles by Selby & Bear (1956) provides additional support for the correctness of the analogue. These particular smooth muscles are known to consist of thin filaments of actin and thick filaments of paramyosin (Hanson & Lowy, 1961). The two non-net reflections were observed by Selby & Bear (1956) in the X-ray diffraction pattern of the translucent parts of the adductor muscle of the mollusc *Venus mercenaria*. The reflections were on the first and second layer lines of an axial periodicity of about 400 Å which did not belong to the moderate-angle series of reflections from actin. It was not known for certain whether these reflections were meridional or not, because they overlapped the central diffuse scatter; however, the second layer line reflection was described as a non-meridional reflection. This diffraction would appear to come from independently diffracting filaments, and could arise in the same manner as the 410 Å meridional reflection of vertebrate muscle and the 410 Å layer line reflections of insect muscle. Since the paramyosin structure has an axial period of 725 Å, it is reasonable to assume that some structural component that gives rise to the non-net reflections of smooth muscle is associated with the actin filaments. This is in accord with the analogue.

The extra reflections from insect, vertebrate striated, and molluscan smooth muscles, which are thought to arise from some component other than the structure of the molecular filaments themselves, have been interpreted in terms of a structural component thought to be related to the cross-bridges seen in the electron microscope. It is encouraging that the extra reflections of different muscles can be observed and attributed to a common origin; nevertheless, it is clear that before a complete description of the large-scale molecular structure of muscle can be given, the present X-ray studies, and also the electron microscope studies, need to be carried a stage further.

The discussion given here, based primarily on X-ray diffraction studies, has directed attention to the lateral register of filaments as an important factor in the way actin and myosin filaments pack together in the myofibril.

I wish to thank Professor J. T. Randall, F.R.S. for the provision of facilities, Dr. K. Machin for supplying the lamellicorn beetles and the Anti-Locust Centre, London, for supplying the desert locusts. I would also like to thank Dr. M. H. F. Wilkins, F.R.S. for valuable suggestions, and Dr. J. Hanson and Dr. H. E. Huxley, F.R.S. for helpful criticism.

REFERENCES

Astbury, W. T. (1947). *Proc. Roy. Soc.* B, **134**, 303.
Boettiger, E. G. (1960). *Ann. Rev. Ent.* **5**, 1.
Cohen, C. & Hanson, J. (1956). *Biochim. biophys. Acta*, **21**, 177.
Draper, M. H. & Hodge, A. J. (1949). *Aust. J. Exp. Biol. Med. Sci.* **27**, 465.
Elliott, G. F. (1960). Ph.D. thesis, London University.
Fischer, E. (1947). *Ann. N.Y. Acad. Sci.* **47**, 783.
Hanson, J. & Lowy, J. (1961). *Proc. Roy. Soc.* B, **154**, 173.
Harris, E. J. (1960). *Transport and Accumulation in Biological Systems*. London: Butterworth Scientific Publications.
Hodge, A. J. (1955). *J. Biophys. Biochem. Cytol.* **1**, 361.
Huxley, H. E. (1953*a*). *Biochim. biophys. Acta*, **12**, 387.
Huxley, H. E. (1953*b*). *Proc. Roy. Soc.* B, **141**, 59.
Huxley, H. E. (1957). *J. Biophys. Biochem. Cytol.* **3**, 631.
Huxley, H. E. & Hanson, J. (1956). *Proceedings of the Conference on Electron Microscopy*. Stockholm: Almqvist and Wiksell.
Huxley, H. E. & Hanson, J. (1960). In *Structure and Function of Muscle*, Vol. 1. New York: Academic Press.
Machin, K. E. & Pringle, J. W. S. (1959). *Proc. Roy. Soc.* B, **151**, 204.
Pringle, J. W. S. (1957). *Insect Flight*. Cambridge University Press.
Selby, C. C. & Bear, R. S. (1956). *J. Biophys. Biochem. Cytol.* **2**, 71.
Spencer, M. & Worthington, C. R. (1960). *Nature*, **187**, 388.
Weis-Fogh, T. (1956). *J. Exp. Biol.* **33**, 668.
Worthington, C. R. (1959). *J. Mol. Biol.* **1**, 398.
Worthington, C. R. (1960). *Acta Cryst.* **13**, 1053.

Appendix

I. Calculation of the $hki0$ reflections; space groups $P6$ and $P6_1$.

The general expressions for A and B reduce to

$$A = 2(\cos 2\pi hx + \cos 2\pi kx + \cos 2\pi ix)$$
$$B = 0$$

Myosin: $x = 0 \quad A = 0$

Actin: $x = \frac{1}{2} \quad A = \Delta_{hk}$

where $\Delta_{hk} = 3$ if h, k even or, otherwise $\Delta_{hk} = -1$.

Cross-bridges: 5 atoms, coordinates $x_1\ x_2\ .\ x_r x_5$.

$$A = 2\sum_r^5 [\cos 2\pi hx_r + \cos 2\pi kx_r + \cos 2\pi ix_r]$$

and

$$A(00\bar{0}0) = +30{\cdot}0$$
$$A(10\bar{1}0) = +3{\cdot}2$$
$$A(11\bar{2}0) = -3{\cdot}4$$
$$A(20\bar{2}0) = +1{\cdot}8 \text{ and so on.}$$

II. Calculation of the $hk\bar{i}1$ reflections

(1) Space group $P6$.

The general expressions for A and B reduce to

$$A = 2 \cos 2\pi lz \,(\cos 2\pi hx + \cos 2\pi kx + \cos 2\pi ix)$$
$$B = 2 \sin 2\pi lz \,(\cos 2\pi hx + \cos 2\pi kx + \cos 2\pi ix)$$

Myosin: $A = B = 0$ as the z coordinate varies continuously from 0 to 1,

Actin: $x = \frac{1}{2}$

$$A = \sum_{\Delta_{hk}} \cos 2\pi lz$$
$$B = \sum_{\Delta_{hk}} \sin 2\pi lz$$

Cross-bridges: $z = 0 \quad B = 0$

$$A(hk\bar{i}0) = A(hk\bar{i}1)$$

(2) Space group $P6_1$.

The general expressions for A and B reduce to

$$A = \sqrt{3}\,[\sin 2\pi ix - \sin 2\pi kx]$$
$$B = 2 \sin 2\pi hx - \sin 2\pi kx - \sin 2\pi ix$$

Myosin: $x = 0 \quad A = B = 0$

Actin: $x = \frac{1}{2} \quad A = B = 0.$

Cross-bridges: 5 atoms, coordinates $x_1x_2 \,.\, x_rx_5$.

$$A = \sqrt{3} \sum_{r}^{5} [\sin 2\pi ix_r - \sin 2\pi kx_r]$$
$$B = \sum_{r}^{5} [2 \sin 2\pi hx_r - \sin 2\pi kx_r - \sin 2\pi ix_r]$$

Write

$$G^2 = A^2 + B^2$$
$$G^2(00\bar{0}0) = 900$$
$$G^2(00\bar{0}1) = 0$$
$$G^2(10\bar{1}1) = 187$$
$$G^2(11\bar{2}1) = 40$$
$$G^2(20\bar{2}1) = 7$$
$$G^2(21\bar{3}1) = 80$$
$$G^2(22\bar{4}1) = 3$$
$$G^2(31\bar{1}1) = 46 \text{ and so on}$$

J. Mol. Biol. (1961) **3**, 634–639

Interaction of DNA and Proflavine: A Small-Angle X-Ray Scattering Study

V. Luzzati, F. Masson

Centre de Recherches sur les Macromolecules, 6 rue Boussingault, Strasbourg, France

AND

L. S. Lerman†

Department of Biophysics, University of Colorado Medical Center, Denver, Colorado, U.S.A.

(*Received 28 April 1961*)

The interaction of DNA and proflavine has been studied in isotropic dilute solutions by small-angle X-ray scattering measurements, performed on an absolute scale. The structure of DNA, and of the DNA-proflavine complex, is rod-like: the mass per unit length of the rods and the radius of gyration around the axis have been measured. The variation of these parameters, as a function of the proflavine concentration, is shown to agree with a mechanism of interaction whereby the proflavine molecules are intercalated between nucleotide-pair layers into the DNA helix.

1. Introduction

Peacocke & Skerrett (1956) have shown that proflavine is bound to DNA by two different mechanisms: a strong, first-order combination which reaches saturation at one proflavine per four or five nucleotides, and a weaker, higher-order process leading to the fixation of one dye molecule per nucleotide. Structural aspects of the first-order combination were investigated by Lerman (1961) by means of measurements of the viscosity and sedimentation of the complex in solution and comparison of the X-ray diffraction patterns of DNA fibers with and without proflavine. It was concluded that proflavine is intercalated between nucleotide-pair layers into the DNA helix by extension and some untwisting of the phosphate-deoxyribose backbone, with both the bases and the dye molecules perpendicular to the helix axis. Other models were considered and discarded as incompatible with either the hydrodynamic or the diffraction experiments, the observations consisting principally of a diminished mass-per-unit length (corresponding to the decreased sedimentation coefficient of the complex) and the retention of the 3·4 Å meridional spacing. Using a worm-like chain model for DNA, the intercalation hypothesis permitted calculation of the intrinsic viscosity of the complex in satisfactory agreement with experiment.

Since we have recently developed, at the Centre de Recherches sur les Macromolecules, a small-angle X-ray scattering technique that determines the linear mass of rod-like molecules in solution, we were able to confirm the diminished mass-per-unit length of the DNA-proflavine complex as compared with DNA alone.

† National Institutes of Health Senior Research Fellow.

2. Methods

DNA was extracted from calf thymus by the usual procedures (Mirsky & Pollister, 1946). The proflavine hemisulphate was the same commercial sample previously used by Lerman (1961). The samples for X-ray experiments were prepared by adding proflavine in 0·1 M-NaCl to DNA dissolved in the same solvent; the solutions were allowed to stand to reach equilibrium for a few days. The DNA concentration was determined by phosphorus analysis of each sample after X-ray measurements. Although the DNA concentrations are much higher than in Lerman's hydrodynamic experiments, we may assume that the general structural features of the combination are not sensitive to DNA concentration. The proflavine concentrations in all of the experiments described here are restricted to the first order, strong-binding region, i.e., less than one dye molecule per four nucleotides. (Peacocke & Skerrett, 1956). When proflavine was added in excess of this amount we observed the formation of two phases: a hard, yellow complex, and a fairly fluid solution; this phase separation has also been described by Peacocke & Skerrett.

The X-ray technique has been described elsewhere (Luzzati, 1960; Luzzati, Nicolaieff & Masson, 1961); we have shown how the experimental curves are normalized, leading to a function $j_n(s)$ that depends only on the structure of the sample. In our notation (Luzzati *et al.*, 1961), s is the diffraction angle ($s = 2 \sin \theta/\lambda$), c_e is the concentration (the ratio of the number of electrons of the sodium salt of DNA to the number of electrons of the solution), ρ_0 is the electron density of the solvent (el. Å^{-3}), ψ is the electronic partial specific volume of NaDNA in water (Å^3 el.$^{-3}$), μ is the linear electron content of NaDNA rods (el. Å^{-1}), R_c is the radius of gyration around the rod axis of the difference between the electron density inside the rod and in the solvent. The four bases are assumed to be in equimolar concentration of DNA. The molecular weight of the tetranucleotide (Na salt) is 1,323; the number of electrons is 680. The amount of proflavine is given by the ratio, β, of the proflavine molecules to the nucleotide pairs.

3. Results and Discussion

For a solution of identical rod-like particles the experimental curves obey the theoretical expression (Luzzati, 1960; Luzzati *et al.*, 1961):

$$j_n(s) = A[\tfrac{1}{2} \exp(-\pi^2 R_c^2 s^2) K_0(\pi^2 R_c^2 s^2)] \quad (1)$$

$$A = c_e \mu (1 - \rho_0 \psi)^2 \quad (2)$$

at least at low concentrations where interactions are negligible.

As a convenient representation of the experimental curves, we have plotted $\log j_n(s)$ as a function of $\log s$; we have discussed the elimination of the interaction effects, and described how A and R_c can be determined. The experimental curves are plotted in Fig. 1. For every value of β we have made two X-ray experiments, at DNA concentrations of about 2 and 4%.

The experimental curves change in a regular manner as the amount of proflavine increases (Fig. 2) but, in all cases, their shape is typical of rod-like particles. For each experimental curve we have drawn (Fig. 1) the theoretical function, equation (1), corresponding to the values of A and R_c that are given in Table 1.

When A is known, μ can be calculated from equation (2) if the values of c_e and ψ for the DNA-proflavine complex are known. Assuming that the specific volumes of DNA and proflavine are independent, the expressions of c_e and ψ for the complex are (where suffix 1 refers to NaDNA, suffix 2 to proflavine):

$$c_e = c_{e_1} + c_{e_2} = c_{e_1}(1 + \alpha) \quad (3)$$

$$\psi = \frac{c_{e_1}\psi_1 + c_{e_2}\psi_2}{c_{e_1} + c_{e_2}} = \frac{\psi_1 + \alpha\psi_2}{1 + \alpha} \quad (4)$$

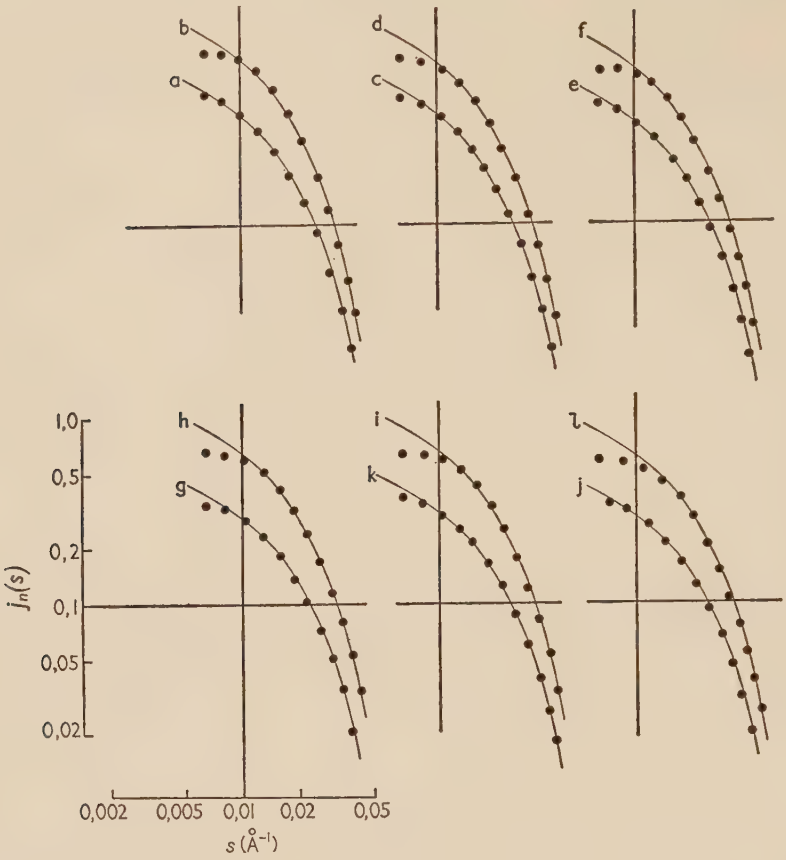


FIG. 1. Experimental curves of $\log j_n(s)$ *versus* $\log s$, and theoretical function, equation (1), corresponding to the experimental curves. DNA and proflavine concentrations, as well as A and R_c, are given in Table 1.

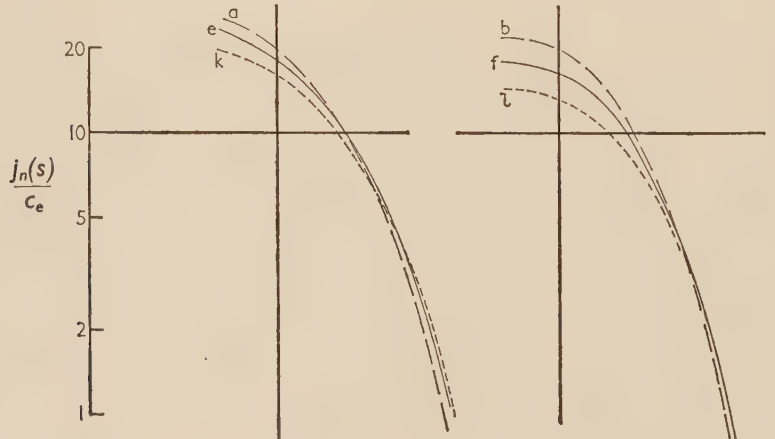


FIG. 2. Two sets of curves of $\log j_n(s)/c_e$ *versus* $\log s$, showing the systematic change of the experimental curves with increasing proflavine concentrations.

α is the ratio of the number of electrons in proflavine free base to the number of electrons in NaDNA ($\alpha = 0{\cdot}324\,\beta$).

By substitution of (3) and (4) into (2), μ becomes:

$$\mu = \frac{A}{c_e(1-\rho_0\psi)^2} = \frac{A}{c_{e_1}(1-\rho_0\psi_1)^2} \cdot \frac{1+\alpha}{\left[1+\frac{\alpha(1-\rho_0\psi_2)}{1-\rho_0\psi_1}\right]^2} \tag{5}$$

Since α is small, (5) can be simplified:

$$\mu = \frac{A}{c_{e_1}(1-\rho_0\psi_1)^2}\left\{1+\alpha\left[1-2\frac{1-\rho_0\psi_2}{1-\rho_0\psi_1}\right]\right\} \tag{6}$$

It appears that the specific volume of proflavine has not been measured, so, by analogy with acridine, whose crystal density is 1·24 g cm^{-3} (Lowde, Phillips & Wood, 1953), we estimate v for proflavine to be roughly 0·80 ($\psi_2 = 2{\cdot}50$ Å^3 el.$^{-1}$). The other parameters of (6) are known: $\rho_0 = 0{\cdot}334$ el. Å^{-3}, $\psi_1 = 1{\cdot}85$ Å^3 el.$^{-1}$ (Luzzati *et al.*, 1961).

Table 1

	β	c_e (%)	R_c (Å)	A	A/c_e	μ (el. Å^{-1})
a	0	2·04	8·4	0·299	14·65	100·4
b	0	3·95	8·4	0·596	15·09	103·2
c	0·103	1·94	8·2	0·290	14·95	102·5
d	0·097	3·88	8·2	0·544	14·01	96·0
e	0·208	1·92	8·0	0·257	13·41	91·9
f	0·202	3·82	8·0	0·488	12·78	87·5
g	0·324	1·89	7·8	0·211	11·17	76·5
h	0·304	3·83	7·8	0·465	12·12	83·0
k	0·420	1·92	7·6	0·218	11·35	77·7
i	0·348	3·83	7.7	0·465	12·12	83·0
j	0·456	1·69	7·5	0·202	11·95	81·9
l	0·510	4·06	7·5	0·426	10·50	71·9

With these numerical values the term inside the brackets in (6) is so very close to unity that it can be neglected. μ then takes the simple form

$$\mu = \frac{A}{c_{e_1}(1-\rho_0\psi_1)^2} \tag{7}$$

The values of μ and R_c are given in Table 1 and are plotted in Fig. 3. The variation of μ and R_c with the amount of proflavine shows that:

(a) the interaction with proflavine modifies the structure of DNA in solution but preserves the rod-like configuration,

(b) the linear mass of the rods decreases, and

(c) the radius of gyration, R_c, decreases with increasing proflavine concentration.

These observations are in qualitative agreement with the model of proflavine molecules intercalating into the DNA helix by extension of the backbone (Lerman, 1961), and are incompatible with any stacking of proflavine molecules on the outside of DNA in the *B* configuration.

The agreement with the intercalated model can be further tested quantitatively. If the distance between planar rings (projected on the helix axis) is the same (3·36 Å)

in DNA and in the complex, and if two sodium ions are displaced by one proflavine base, μ is

$$\mu = \frac{340 + \beta(110 - 22)}{3{\cdot}36\ (1 + \beta)} = 101 - 75\frac{\beta}{1 + \beta} \tag{8}$$

The straight line representing μ as a function of $\beta/(1+\beta)$, (8), is drawn in Fig. 3, where the experimental values of μ are plotted; the agreement is as satisfactory as can be expected if the experimental errors are taken into account (the average discrepancy of the experimental points from the theoretical line is about 4%, i.e. similar to previous work on DNA (Luzzati *et al.*, 1961).

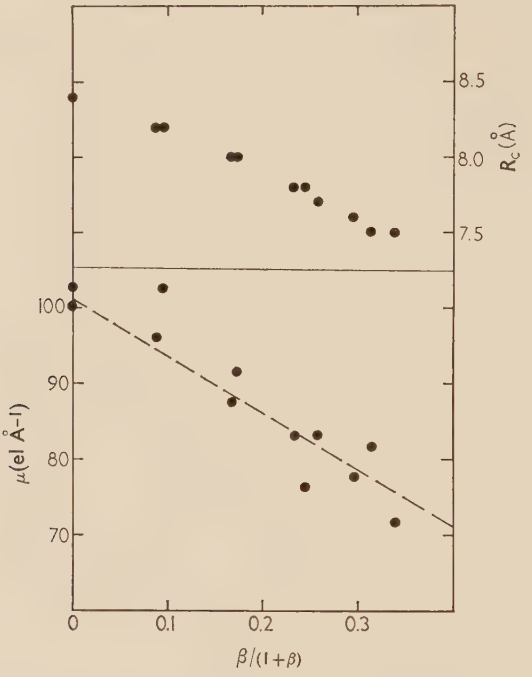


FIG. 3. Experimental values of μ and R_c, as a function of $\frac{\beta}{1+\beta}$ (see Table 1). The straight line corresponds to the intercalated model.

In the previous investigation, two additional models for the mode of binding were considered. One of these, a hypothetical configuration in which proflavine is inserted between nucleotide-pair layers by outward separation and displacement of a pair, can be discarded because it corresponds to an increased linear mass. However, the other proposal, in which the external attachment of the dye induces tilting of the base pairs to a configuration analagous to that of stretched DNA fibers (Wilkins, Gosling & Seeds, 1951), cannot be distinguished from intercalation by the linear mass measurement alone. In stretched DNA, the 50% extension of the fiber must correspond to a 1/3 diminution of the linear mass. If it is supposed that the attachment of each proflavine induces on average the tilting of 2·3 base-pairs, at the same time displacing two sodium ions (see above), the linear mass becomes

$$\mu = 101 - 52\beta \tag{9}$$

This would result in a curve slightly convex upward in Fig. 3, matching equation (8)

at the ordinate and at $\beta/(1+\beta) = 0{\cdot}3$, and fitting the experimental points at least as well. While this model appears somewhat forced, in that a process of this kind would be expected to be higher than first order in its dependence on bound dye, its exclusion depends on the determination of the orientation of the base pairs. It was shown (Lerman, 1961) that fibers of DNA-proflavine retain the 3·4 Å spacing characteristic of bases perpendicular to the helix axis and fail to present the longer spacing expected for tilted bases. There is no reason to expect the configuration to be different in solution.

The variation of R_c shows that on combination with proflavine the DNA becomes more compact. It may be supposed that the change is due to contributions from several factors—the displacement of sodium ions, the smaller radius of the intercalated proflavine, and perhaps a contraction of the phosphate radii toward the proflavine. There may also be a contribution from the shift of the center of gyration from the original helix axis to near the center of mass of the base pairs, owing to the four times longer pitch of the helix containing proflavine; it has been suggested that the longer pitch is responsible for the slightly closer packing in the fiber.

REFERENCES

Lerman, L. S. (1961). *J. Mol. Biol.* **3,** 18.
Lowde, R. D., Phillips, D. C. & Wood, R. C. (1953). *Acta Cryst.* **6,** 553.
Luzzati, V. (1960). *Acta Cryst.* **13,** 939.
Luzzati, V., Nicolaieff, A. & Masson, F. (1961). *J. Mol. Biol.* **3,** 185.
Mirsky, A. E. & Pollister, A. W. (1946). *J. Gen. Physiol.* **36,** 117.
Peacocke, A. R. & Skerrett, J. N. H. (1956). *Trans. Faraday Soc.* **52,** 261.
Wilkins, M. H. F., Gosling, R. G. & Seeds, W. E. (1951). *Nature,* **167,** 759.

J. Mol. Biol. (1961) **3**, 640–647

The Molecular Configuration of α-Keratin

R. D. B. Fraser and T. P. MacRae

Division of Protein Chemistry, Wool Research Laboratories, C.S.I.R.O. Parkville N2, (Melbourne), Victoria, Australia

(*Received 2 May 1961*)

The coiled-coil configurations of the α-helix which have been suggested by Crick and Pauling & Corey as models for the structure of α-keratin have been tested by comparing their optical transforms with the observed X-ray diffraction pattern. The observed meridional low-angle reflections may be indexed as orders of an axial repeat of 197 Å and this has in the past been associated with the pitch of the coiled-coil model. Evidence is presented which suggests that coiled coils if present do not persist over sufficient lengths to justify this conclusion. The coiled-coil projections used to obtain the optical transforms were therefore restricted to short lengths (~ 60 Å) and the results obtained with the 3-strand parallel-chain rope show good agreement with the observed wide-angle X-ray pattern except in the 1 Å region of the meridian. It is shown that a modification of the coiled-coil model in which the distortion associated with coiling is localised at particular residues rather than being uniformly distributed gives a better agreement in the 1 Å region.

1. Introduction

It is now generally accepted that the polypeptide chain configuration in α-keratin is closely related to that of the α-helix described by Pauling, Corey & Branson (1951). However, in its simplest form the α-helix does not account for the simultaneous appearance of the two prominent meridional arcs of spacing 5·1 and 1·49 Å observed in the wide-angle X-ray diffraction pattern of α-keratin. Pauling & Corey (1953) and Crick (1952, 1953) suggested independently that these reflections could be explained by supposing that the α-helices were systematically distorted into "coiled coils," either by the occurrence of a repeating sequence of amino acid residues or to achieve optimum packing. Detailed calculations by Crick (1953) and Lang (1956) showed that the 5·1 and 1·49 Å meridional reflections were satisfactorily accounted for by a model in which optimum packing of three α-helices was obtained by coiling them around a common axis to form a 3-strand rope. Confirmation of these findings by means of optical diffraction studies was reported recently in a preliminary note (Fraser & MacRae, 1961) and in the present communication the elaboration and extension of these studies is described.

2. Experimental

X-ray diffraction diagrams were obtained from the quill tips of the North American porcupine (*Erethizon dorsatus*) and South American porcupine (*Coendou paraguayensis*) using CuKα radiation and recorded either on a 3 cm cylindrical camera or a 10 cm

low-angle pinhole camera capable of resolving 200 Å. Optical transforms of various model structures were obtained with a Pullin optical diffractometer equipped with 4·5 in diameter lenses. The atoms were represented by 0·3 mm diameter holes punched in heavily exposed X-ray film using a scale of 0·6 mm to 1 Å. Owing to the very great complexity of the resulting masks and the semi-quantitative nature of the investigation it was not considered profitable at this stage to differentiate between the scattering powers of the CH_2, CH, NH and O atomic groupings, although this refinement may be necessary in subsequent investigations. In some instances the fine detail in the optical transforms was smeared by oscillating the mask through a small angle during photographic recording to produce a more realistic representation of the state of orientation in the natural material.

In the case of the coiled-coil models the transforms were sampled at 180° intervals; with the segmented rope model the stationary transform was similarly sampled whilst the oscillated transform was sampled at 90° intervals. Thus the illustrations in Plates I and II were obtained by an intensity addition of the diffraction pattern from the mask and after inversion about the fibre axis, whilst that of Plate III(d) was obtained from two masks together with their inversions.

A refinement which might be justified in any further studies is a more complete sampling of the transforms as recommended by Stokes (1955).

3. The Scope of the Investigation

The diffraction pattern yielded by porcupine quill tip is extremely complex (MacArthur, 1943) and the more prominent features observed in the present investigation are collected together in the composite diagram of Plate I(a). The meridional and near-meridional low-angle reflections shown in Plate IV(a) may be indexed as orders of an axial repeat of 197 Å and Crick (1953), Pauling & Corey (1953) and Lang (1956) have all associated this length with the pitch of the coiled-coil model. There is, however, a considerable body of evidence which will be discussed in detail in Section 6 suggesting that if coiled coils are present they do not persist with geometrical regularity over a sufficient number of turns to justify such a conclusion. Accordingly the present investigation has been restricted to very much shorter lengths of coiled coil than were envisaged in the models described by Crick (1953) and Pauling & Corey (1953). A convenient choice from the experimental point of view is 62 Å, the repeating unit of the 3-strand rope. Even so the labour involved in computing the coordinates of the 630 atoms in this unit is very considerable and the method used in the preliminary investigation has since been found to be unreliable. Consequently the transform in Plate I(a) of our preliminary publication (Fraser & MacRae, 1961) differs in some respects from the corresponding transform in Plate II(b) of this paper. In all the projections calculated for the present work the method given by Crick (1953) was used.

4. The Coiled-Coil Models

Details of the coiled-coil models examined are as follows, the notation being that used by Crick (1953).

(a) *Single coiled coil*

Turns of major helix	$N_0 = 1$
Turns of minor helix	$N_1 = 36$
Atoms in set	$M = 126$
Repeat distance	$C = 186$ Å
Radius of major helix	$r_0 = 5{\cdot}2$ Å

The projection used was that of a 62 Å length containing sets of 5 atoms corresponding to the C′, O, N, C-α and C-β atoms of the peptide group. The sense of the major helix was left-handed, the minor helix right-handed and the configuration about the α-carbon atoms was that corresponding to L-amino acids.

(b) *Parallel-chain 2-strand rope*

Two coiled coils having the parameters of the single coiled coil given above were combined in the manner suggested in Fig. 4 of Crick's paper so that the two chains were related by a dyad axis. The repeat of structure was therefore reduced to 93 Å.

(c) *Anti-parallel chain 2-strand rope*

By choosing a suitable origin and plane for the projection it was possible to obtain the coordinates of a 2-strand rope in which the chain sense (NH . CHR . CO $\longrightarrow$) was oppositely directed in the two helices, without further calculation. In this model the β-carbon atoms of the two chains occupied the same positions as in the parallel-chain 2-strand rope so that "knob-hole" packing was still achieved but the remainder of the atoms in one chain were no longer related by a dyad axis with the corresponding atoms in the second chain. The repeat of structure in this case was 186 Å.

(d) *Parallel-chain 3-strand rope*

The parameters of each coiled coil in the 3-strand rope were identical with those of the single coiled coil except that r_0 was increased to 5·5 Å; this necessitated recalculation of all the atomic positions. The orientation of the β-carbon atoms was chosen to conform with the packing arrangement shown in Fig. 6 of Crick's paper. Thus the 3 chains were related by a triad axis and the repeat of structure was reduced to 62 Å.

(e) *Two parallel, one anti-parallel chain 3-strand rope*

By a suitable choice of origin and projection plane the coordinates of this model were obtained without further calculation. The positions occupied by the β-carbon atoms in the anti-parallel chain were the same as those in the parallel-chain 3-strand rope so that "knob-hole" packing was retained but the remainder of the atoms were no longer related by a triad axis; the period therefore was 186 Å. The transform of the one-parallel, two anti-parallel chain 3-strand rope is of course obtained simply by rotating the projection through 180° about the projection axis.

5. The Optical Transforms of the Coiled-Coil Models

The optical transforms of the coiled-coil models are shown in Plates I and II and it will be seen that the meridional reflections are restricted to spacings which are submultiples of 10·33 Å. In every case the 7th order of 10·33 Å (1·475 Å) is prominent, thus accounting satisfactorily for the observed reflection at 1·485 Å. A second feature common to all the models is the splitting of the lateral interference about 10 Å into two near-equatorial components. Although we have not succeeded in completely resolving the group of reflections centred at 9·8 Å in the observed diffraction pattern there is clear evidence (Plate I(a)) that it consists of both equatorial and near-equatorial maxima. The most clear cut example is the resolution of a pair of near equatorial reflections with coordinates $\xi = 0{\cdot}08$ Å^{-1} and $\zeta = 0{\cdot}0143$ Å^{-1}.

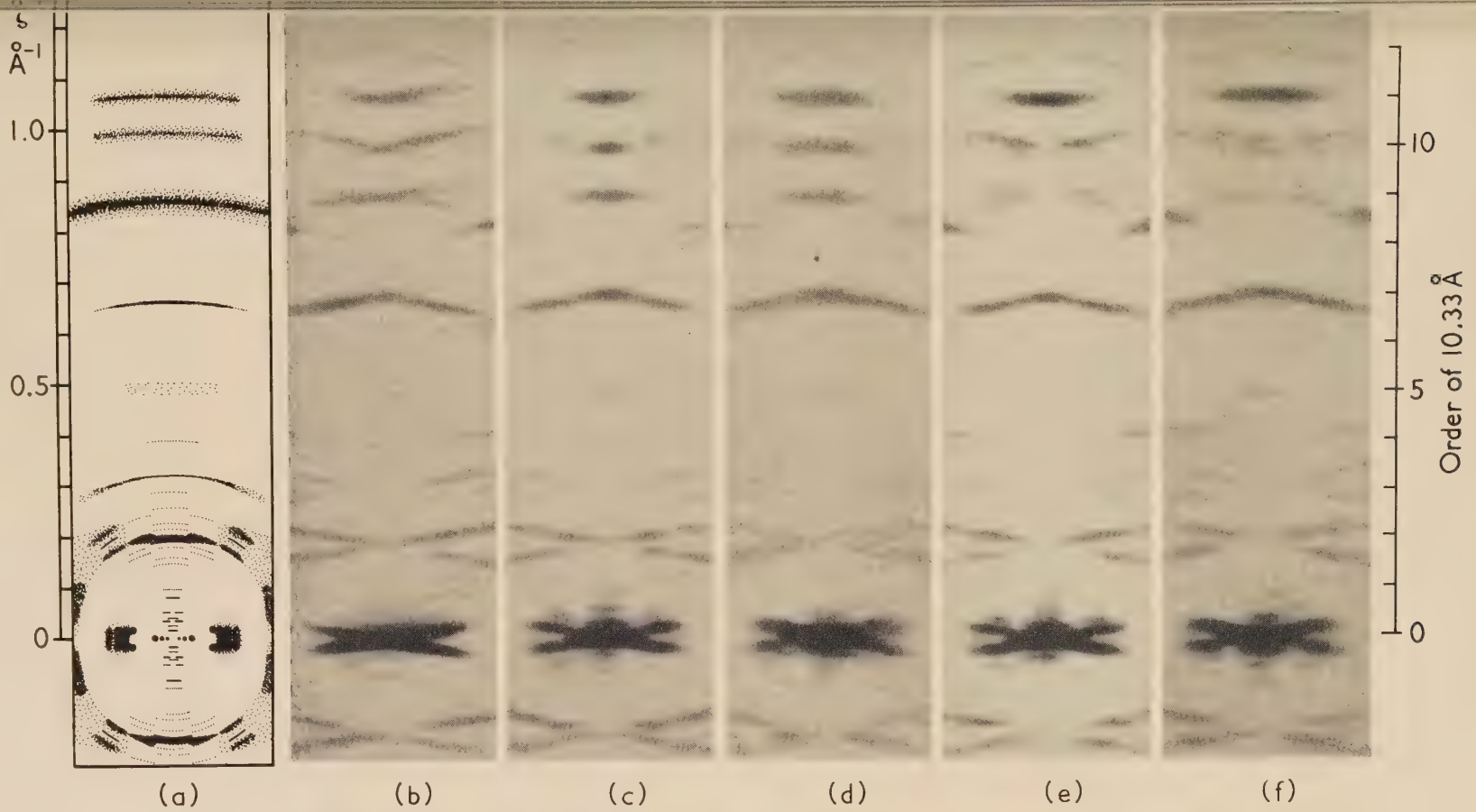


PLATE I.

(a) Composite diagram illustrating the main features observed in the X-ray diffraction diagram of porcupine quill tip; (b) optical diffraction pattern obtained from a projection of a single coiled coil; (c) the 2-strand parallel chain rope; (d), (c) oscillated through a small angle during exposure; (e) the 2-strand antiparallel chain rope; (f), (d) oscillated.

[*To face page* 642

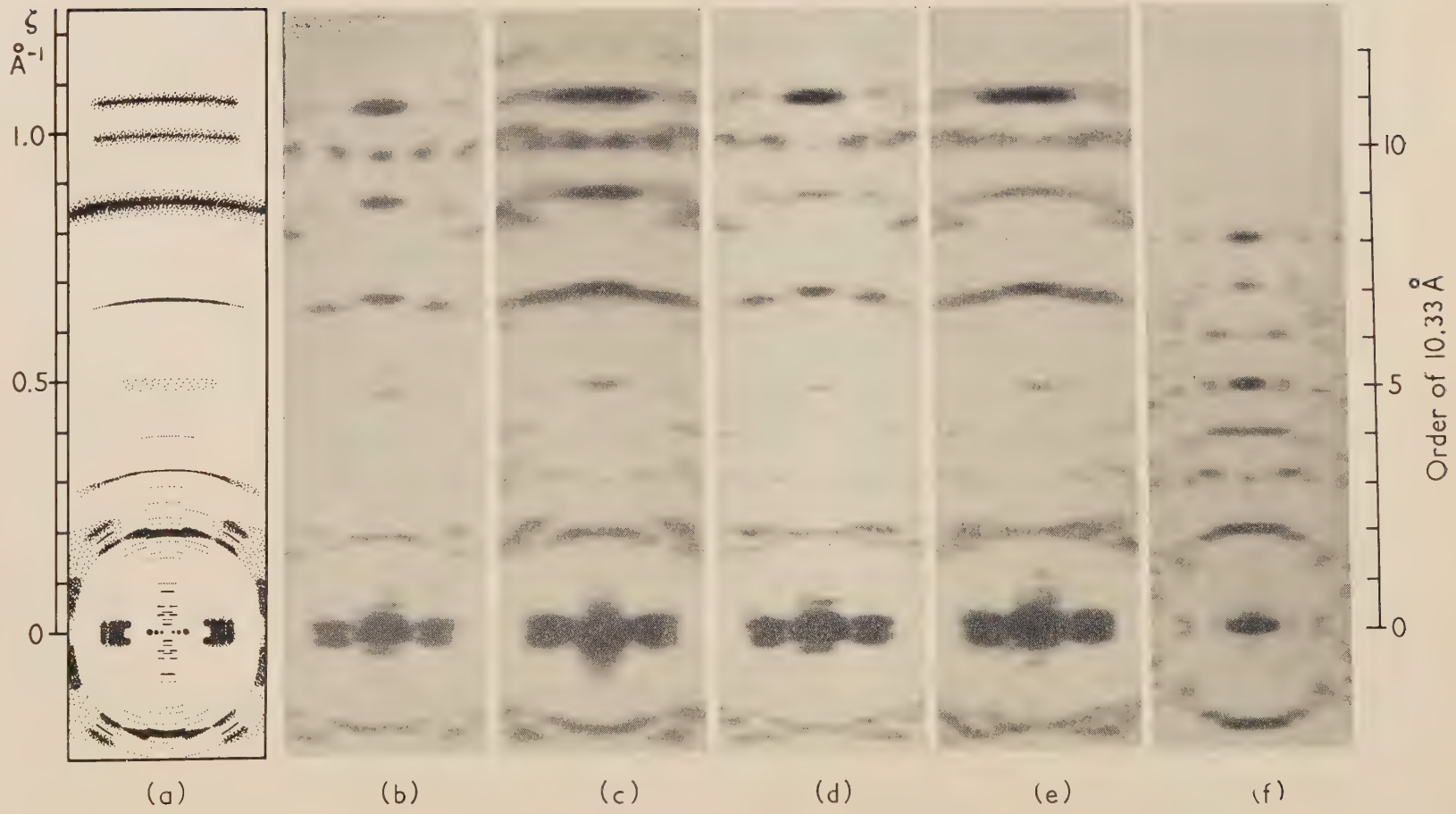


PLATE II.

(a) Observed X-ray pattern; (b) 3-strand parallel-chain rope; (c), (b) oscillated; (d) 3-strand rope, two parallel one anti-parallel chain; (e), (d)

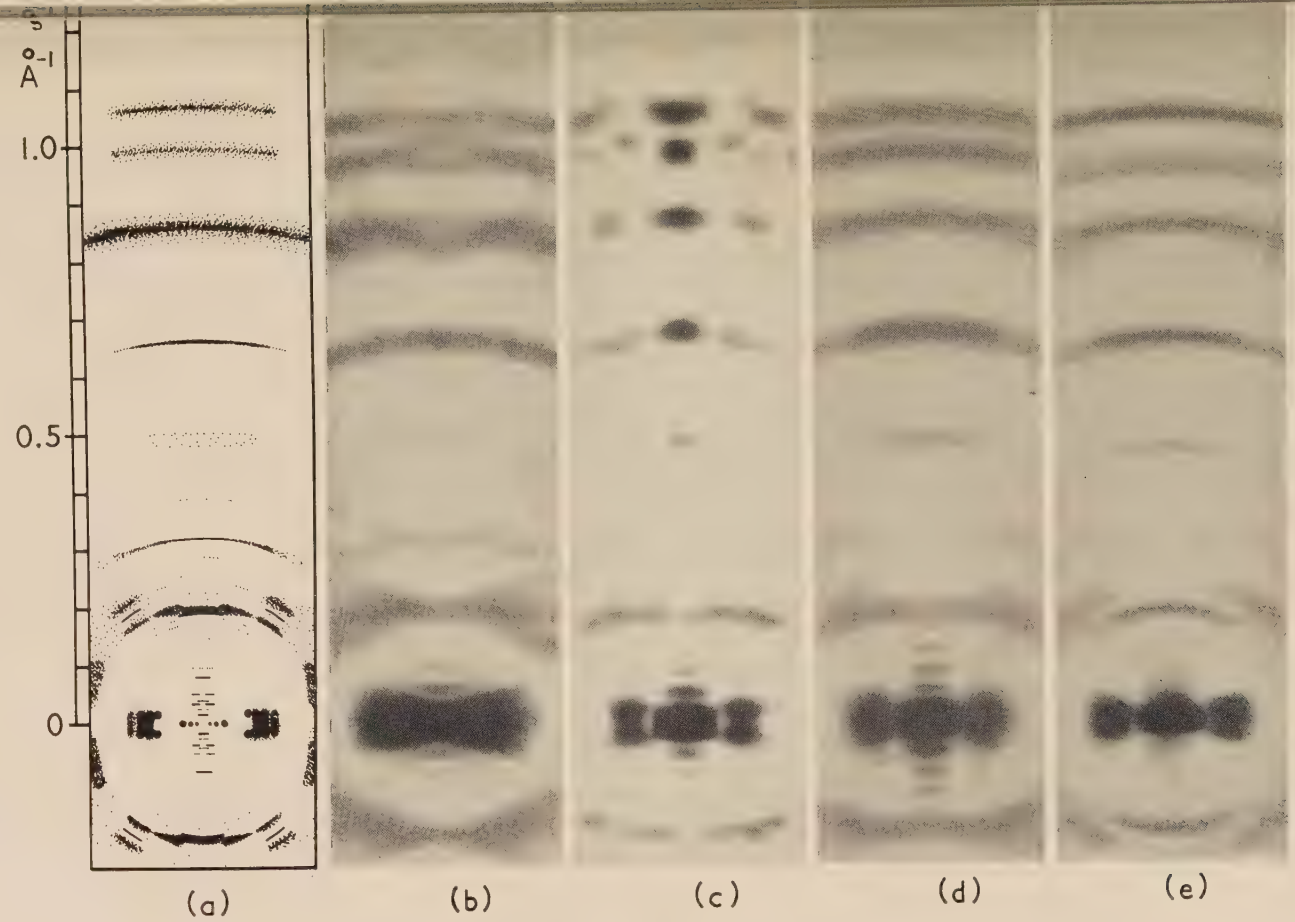
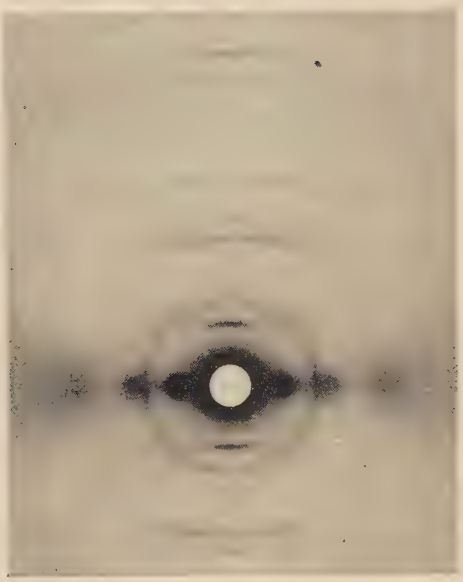


PLATE III.

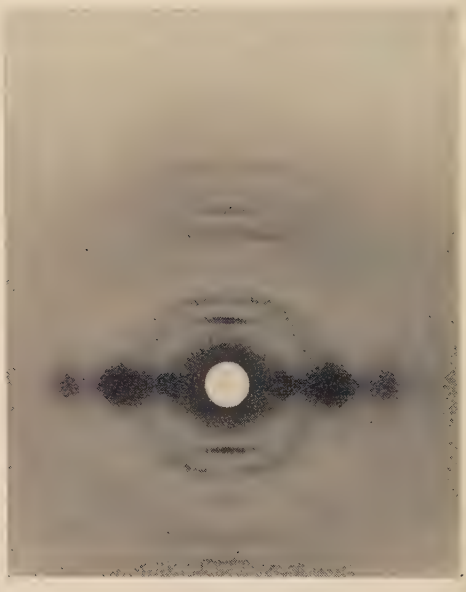
(a) Observed X-ray pattern; (b) intensity addition of the optical diffraction patterns of thirteen 27 Å lengths of α-helix with axial directions distributed over a cone of semi-angle 20°; (c) intensity addition of one aspect of a group of three α-helices inclined at 10° to the fibre axis and its inversion; (d) intensity addition of 2 aspects of (c) (oscillated ± 10°); (e) 3-strand parallel chain rope (oscillated ± 10°).



(a)

(b)

(c)



(d)

PLATE IV. Low-angle X-ray diffraction patterns obtained from dry North American porcupine quill tip.

(a) Untreated.
(b) Stained with 1% w/v osmium tetroxide for 3 days at 20°C.
(c) Treated with 0·3 N-nitric acid for 3 days at 68°C (*cf.* Kratky *et al.*, 1955).
(d) Stained with 1% w/v silver nitrate for 7 days at 20°C.

As suggested by Crick (1953) the transforms do not have sufficient intensity on the meridian in the region of the 2nd order of 10·33 Å to account for the very strong observed reflection at 5·14 Å. However the transform shown in Plate II(f) of a 3-strand rope with a single atom set with $r_1 = 4$ Å clearly demonstrates the important contribution which the side-chains could make to this reflection. The possible contribution of the side-chains to the near equatorial reflections is also evident in this transform.

In many respects then coiled-coil models provide a plausible explanation of certain details of the observed diffraction pattern. The only serious discrepancy appears to be in the very wide-angle meridional diffraction where the reflections of the coiled coils are restricted to spacings which are sub-multiples of 10·33 Å whilst the observed pattern does not show this feature. Some account must be taken however of the effects of disorientation which smear near-meridional reflections across the axis of the observed pattern, and it will be noted that in Plate II(e) a slight improvement in agreement was obtained by oscillating the mask. The additional possibility that the lengths of coiled coil may be appreciably less than 62 Å was also investigated but no significant improvement was noted.

Table 1

Materials	Axial period (Å)	Some prominent meridional or very nearly meridional reflections (Å)					
S. American porcupine quill	197			12·3 [g]		24·2 [g]	66·6 [g]
						27·3 [g]	
African porcupine quill	198	1·486 [b]	5·11 [b]	12·4 [e]		24·5 [d]	66·3 [e]
		1·485 [c]	5·14 [e]			27·4 [d]	
Light meromyosin	430				15·9 [f]	26·8 [f]	70·8 [f]
Proteus vulgaris flagella	410	1·49 [a]	5·11 [a]	12·8 [a]	15·8 [a]	25·2 [a]	—
Bacillus subtilis flagella	410		5·11 [a]	12·5 [a]	16·0 [a]	25·7 [a]	—

[a] Astbury, Beighton & Weibull (1955)
[b] Astbury & Haggith (1953)
[c] Bamford, Elliott & Hanby (1956)
[d] Bear (1944)
[e] MacArthur (1943)
[f] Szent-Györgyi, Cohen & Philpott (1960)
[g] Present work

A further minor difficulty in this respect is that the ratio of the 5·1 and 1·49 Å reflections should be exactly 3·5, as they are the 2nd and 7th orders of 10·33 Å. The best estimate of the exact spacing of the 7th order (Table 1) is 1·485 Å so that the 2nd order should be 5·20 Å. This is certainly outside the experimental error for the value of 5·14 Å reported for this reflection, but again it is possible that the fine structure in this region has not as yet been completely resolved.

6. The Segmented Rope Model

In a preliminary note (Fraser & MacRae, 1961) we suggested that instead of the uniformly coiled-coil model a type of structure which we designated as a segmented rope might give a better agreement with the observed meridional diffraction pattern. This suggestion was based on the fact that the stability of the α-helix in synthetic polypeptides depends very markedly on the nature of the residue, and we visualized the distortion required to achieve "knob-hole" packing as being concentrated at

residues which do not favour the α-helix, e.g. proline, cystine, tyrosine, serine, valine, phenylalanine and aspartic acid, rather than uniformly distributed over all residues as in the coiled-coil models. Reference to Table 2 shows that appreciable proportions of these residues are present in the fibrillar component of α-keratin. The segmented

TABLE 2

The composition of the low-sulphur S-carboxymethylkeratin extract from merino wool (adapted from Crewther & Dowling, 1960)

Residue in parent protein	(a) μmoles/g	(b) Number per 100	(c) Average axial separation (Å)	(d) Number per 126
Alanine	588	7	23	8
Arginine	681	8	19	9
Aspartic Acid	810	9	16	11
½ Cystine	446	5	30	6
Glutamic Acid	1248	14	11	17
Glycine	602	7	22	8
Histidine	65	1	211	1
Isoleucine	331	4	40	4
Leucine	867	10	15	12
Lysine	356	4	37	5
Phenylalanine	249	3	53	3
Proline	307	3	44	4
Serine	723	8	19	10
Threonine	416	5	32	6
Tyrosine	314	4	42	4
Valine	488	5	27	7
Amide	925	10	14	12

(a) Content as μmoles/g of S-carboxymethyl protein.
(b) Number per 100, assuming that residues not accounted for in the analysis have a residue weight equal to the mean (111) of those recovered.
(c) The average distance between residues of a particular type in the direction of the fibre axis.
(d) The number of residues of a particular type which on average occur in 186 Å, that is, one period of a single coiled coil.

cables so derived would be expected to have an average pitch of about the same length as the coiled coils, and the lengths of the segments of straight α-helix would not be expected to be much in excess of 20 Å owing to the rapid departure from "knob-hole" packing. In order to test this suggestion masks were prepared as follows:

(a) *Disoriented α-helices*

The transform shown in Plate III(b) was obtained by an intensity addition of thirteen 27 Å segments of straight α-helix with axial directions uniformly distributed over the range 0 to 20°. The sense of the α-helices was right handed and a β-carbon atom in the L-configuration was included. It will be seen that the 1·49 and 5·1 Å meridional reflections are satisfactorily accounted for and that the wide-angle diffraction corresponds more closely with the observed pattern than do the coiled-coil models.

(b) *Segmented rope section*

The simple disoriented α-helix model takes no account of the external interference between the α-helices in a segmented cable, and this was introduced by combining three 27 Å lengths of α-helix so that they were related by a triad axis, each helix being inclined at 10° to this axis. The configuration at the centre was arranged to correspond with the packing described earlier for the 3-strand rope.

The diffraction pattern obtained from this segmented rope section, shown in Plate III(c), is necessarily more diffuse than those of Plates I and II owing to the shorter length used, but it will be noted that the 10 Å equatorial reflection is still resolved into two near equatorial components and that the spacings of the meridional reflections are closer to the observed than is the case for the coiled-coil models. This is surprising in view of the widespread belief that the "α"-pattern cannot be explained by tilted α-helices.

It was found that by oscillating the mask of Plate III(c) through $\pm$ 10° during exposure a very close agreement with the appearance of the observed meridional pattern could be obtained, as shown in Plate II(d), although the resolution of the 10 Å equatorial reflection is eliminated by this procedure. A similar improvement was also obtained with the 3-strand parallel chain rope shown in Plate III(e). The measure of agreement between Plates III(a), (d) and (e) provides strong presumptive evidence that rope-like structures are the basis of the molecular configuration in α-keratin as well as in the other members of the *k-m-e-f* group of fibrous proteins, which all exhibit similar wide-angle diffraction diagrams.

Although we have not as yet tested the possibility in detail, it would appear that some modification of the coiled-coil model is required to improve the agreement with the observed pattern, and the evidence from the idealised segmented rope model suggests that a localisation of the distortion at specific residues, resulting in the frequent occurrence of short lengths of straight α-helix, is a probable one.

Of the models covered by the present investigation the 3-strand parallel chain segmented rope appears to give the best agreement with the observed pattern, but it is quite possible that the 7-strand cable proposed by Pauling & Corey (1953) would agree equally well. Unfortunately the atom population of this model in projection is too dense for our existing technique and equipment and we have not been able to investigate this structure.

7. Macromolecular Structure

Although rope-like assemblies of α-helices appear from our optical diffractometer studies to account satisfactorily for the main features of the wide-angle diffraction pattern in α-keratin they do not, in their simplest form, account for the rich detail at low angles or the series of sharp faint meridional reflections which extends well beyond the 5·1 Å arc. In this concluding section the possible relation of rope structures to these features is discussed.

(a) *Equatorial pattern*

The low-angle equatorial diffraction pattern of α-keratin (Plate IV(a)) is dominated by three maxima of spacing 77, 41 and 26 Å which, as we have shown elsewhere (Fraser & MacRae, 1958; Fraser, MacRae & Rogers 1959, 1960), result from the

pseudo-crystalline packing of rod-like bodies (microfibrils) about 70 Å in diameter. The microfibrils are embedded in a cystine-rich matrix, and when porcupine quill tip is stained with osmium tetroxide the preferential uptake of osmium in the matrix leads to a very considerable enhancement of the intensity of the low-angle equatorial maxima (Plate IV(b)). In contrast the wide-angle pattern is not noticeably altered. Thus we may conclude that the rope structures are localised in the microfibrils.

The cross-sectional area of each microfibril is such that it could accommodate a maximum of about fifteen 3-strand ropes or seven 7-strand cables; however, no details of the packing arrangement are known. There is, however, clear evidence of sub-structure in electron micrographs obtained by Rogers (1959). Although imperfectly resolved there is a suggestion that each microfibril is composed of a small number of sub-fibrils, and these may well be 7-strand cables or 3-strand ropes.

The lateral fine structure in the 9·8 Å group of reflections (Plate III(a)) which according to Astbury (1956) can be resolved into maxima of spacing 12·2, 11·2, 10·4, 9·2 and 7·8 Å is too rapid an oscillation in the transform to be associated with rope structures only 20 to 25 Å in diameter. It is more likely to be associated with their mode of packing in the microfibril or to the microfibril lattice (Fraser & MacRae, 1958).

(b) *Meridional pattern*

The series of meridional and near-meridional reflections in porcupine quill tip extending from low to moderately wide angles can be indexed on the basis of a 197 Å axial period and this has been associated, as mentioned earlier, with the pitch of a coiled coil. There are several reasons however for doubting this despite the numerological coincidence. Firstly the axial periods of several fibrous proteins which give typical "α" diagrams (and therefore presumably have rope structures) have been collected in Table 1 and it will be seen that α-keratin is the only case in which the period is similar to that of a coiled coil.

A second difficulty is illustrated in Plate IV where significant variations in the intensities of the meridional reflections have been induced by chemical degradation of certain residues and by the incorporation of a heavy metal stain. In neither case have new meridional reflections which do not index on 197 Å been introduced. The most likely explanation is that 197 Å is the axial projection of the "chemical" repeat of structure. In Plate IV(d) there is, incidentally, an enhancement of a pair of off-meridional arcs barely visible in Plate IV(a) which fall between the 3rd and 4th layer lines. Thus the true axial repeat of structure may be close to that of the other proteins listed in Table 1.

The remaining doubt concerns the amino acid analysis of the "low-sulphur" fibrous component given in Table 2. It has proved possible to isolate protein fractions from α-keratin without significant peptide bond hydrolysis and to separate them into fractions relatively low and high in cystine. The minimum molecular weight of the "low-sulphur" protein is about 8,000 (Harrap, 1956) and the analysis shows significant amounts of proline. In view of this evidence it seems unlikely that a coiled coil structure could persist with sufficient geometrical regularity, and over a sufficient number of turns, to produce layer lines associated with the pitch. Even then it would be necessary to assume that the chemical unit also had precisely this length in view of the results shown in Plate IV. A possible explanation is, of course, that neighbouring ropes have a very long true periodicity but are staggered by a distance equal to the pitch of the cable thus giving an axial period of 197 Å in projection.

Whilst there is insufficient information to draw any definite conclusions about the origin of the axial periods in α-keratin and other members of the *k-m-e-f* group the weight of evidence would seem to favour the view that they are associated with the lengths of the sub-unit which polymerise to give the fibrous elements during biosynthesis of the fibre (Kendrew, 1954).

A point of interest which emerges from Table 2 is the possibility, which Astbury has frequently commented on, that the macromolecular structures of the *k-m-e-f* group are similar. The strong reflection in α-keratin at 66 Å has been suggested as corresponding to a particle length (Lindley, 1957) and a similar correspondence to the 70 Å reflection in light meromyosin is possible (Szent-Györgyi, Cohen & Philpott, 1960). A second common feature is a pair of reflections of spacings about 25 and 12·5 Å, which are of great interest in that they are subject to considerable variation in intensity when heavy metal stains and chemical degradation are employed to modify the diffraction pattern. They are thus associated with the distribution of reactive sites and we have speculated (Fraser & MacRae, 1961) that this periodicity may be correlated with the segment length in the segmented rope.

REFERENCES

Astbury, W. T. (1956). *Proc. Int. Wool Text. Res. Conf.* B, 202.
Astbury, W. T., Beighton, E. & Weibull, C. (1955). *Symp. Soc. Exp. Biol.* **9**, 282.
Astbury, W. T. & Haggith, J. W. (1953). *Biochim. biophys. Acta*, **10**, 483.
Bamford, C. H., Elliott, A. & Hanby, W. E. (1956). *Synthetic Polypeptides*. New York: Academic Press.
Bear, R. S. (1944). *J Amer. Chem. Soc.* **66**, 2043.
Crewther, W. G. & Dowling, L. M. (1960). *J. Text. Inst.* **51**, T775.
Crick, F. H. C. (1952). *Nature*, **170**, 882.
Crick, F. H. C. (1953). *Acta Cryst.* **6**, 689.
Fraser, R. D. B. & MacRae, T. P. (1958). *Biochim. biophys. Acta*, **29**, 229.
Fraser, R. D. B. & MacRae, T. P. (1961). *Nature*, **189**, 572.
Fraser, R. D. B., MacRae, T. P. & Rogers, G. E. (1959). *Nature*, **183**, 592.
Fraser, R. D. B., MacRae, T. P. & Rogers, G. E. (1960). *J. Text. Inst.* **51**, T497.
Harrap, B. S. (1956). *Proc. Int. Wool Text. Res. Conf.* B, 86.
Kendrew, J. C. (1954). *The Proteins*, Vol. 2B, p. 845. New York: Academic Press.
Kratky, O., Sekora, A., Zahn, H. & Fritze, E. R. (1955). *Z. Naturf.* **106**, 68.
Lang, A. R. (1956). *Acta Cryst.* **9**, 436.
Lindley, H. (1957). *Text. Res. J.* **27**, 690.
MacArthur, I. (1943). *Nature*, **152**, 38.
Pauling, L. & Corey, R. B. (1953). *Nature*, **171**, 59.
Pauling, L., Corey, R. B. & Branson, H. R. (1951). *Proc. Nat. Acad. Sci., Wash.* **37**, 205.
Rogers, G. E. (1959). *Ann. N.Y. Acad. Sci.* **83**, 378.
Stokes, A. R. (1955). *Progress in Biophysics*, Vol. 5, p. 140. London: Pergamon Press.
Szent-Györgyi, A. G., Cohen, C. & Philpott, D. E. (1960). *J. Mol. Biol.* **2**, 133.

J. Mol. Biol. (1961) **3**, 648–667

Molecular Interpretation of the Elasticity of Resilin, a Rubber-like Protein

Torkel Weis-Fogh

Zoophysiological Laboratory B, University of Copenhagen, Denmark and Department of Zoology, Cambridge University, England

(*Received 20 May 1961*)

Resilin from the elastic tendon of dragonflies (*Odonata*) was analysed mechanically and optically over a large range of strain and swelling, both in compression and in extension and at hydrogen-ion concentrations from pH 1·8 to 12·3. Under all conditions, the protein behaved as a typical rubber. A detailed comparison between the results and the kinetic theories for short-chain rubber networks showed that the agreement between experiment and theory was almost perfect, with respect to both mechanical and photoelastic properties.

At room temperature and neutral pH, the average modulus G is 6·4 kg cm^{-2} (corrected for swelling), the breaking strength 30 to 40 kg per cm^2 unstrained swollen area, and the length at breaking point about 3 times the unstrained length. Unstrained resilin is isotropic but becomes birefringent on deformation, positive in the direction of extension. The birefringence increases in a non-linear manner with the stress, the stress-optical ratio C' being independent of the degree of swelling and of pH, while it is inversely proportional to the absolute temperature. At room temperature, C' was $1{\cdot}3 \times 10^{-4}$ cm^2 kg^{-1}.

It is concluded that resilin consists of a three-dimensional network of long polypeptide chains which are randomly coiled under all conditions and thermally agitated when the protein is swollen. There are about 60 amino-acid residues between two junction points and at least half of the junction points consist of stable chemical cross-linkages, i.e. one interchain cross-link per 60 to 120 residues. The primary structure of resilin must be such that the secondary structure is so unstable as to be practically absent. On the other hand, the chemical cross-links represent a tertiary structure which extends in three dimensions. The possible importance of such networks is briefly discussed in relation to cytology and muscle thermodynamics.

1. Introduction

Resilin is a newly-discovered structural protein from arthropod cuticle (Weis-Fogh, 1960) which has a unique amino-acid composition (Bailey & Weis-Fogh, 1961) and which exhibits almost ideal rubber-like properties as far as thermoelasticity, deformability and elastic recovery are concerned (Weis-Fogh, 1961). In contrast to other elastic structural proteins it should therefore be possible to compare resilin in detail with the ideal networks described by the kinetic theories of rubber elasticity. In this way valuable knowledge may be obtained about a three-dimensional network of polypeptide chains in terms of average length of chains between junction points, the number of junction points, and of the flexibility of the thermally agitated chains. Moreover, the effect of swelling can be studied since resilin swells isotropically in

water to an extent determined mainly by the hydrogen-ion concentration (Weis-Fogh, 1960). Since unstrained resilin is optically isotropic at all degrees of swelling a study of the photoelastic behaviour may serve as a control for conclusions based upon purely mechanical analyses and may also offer valuable quantitative information, hitherto lacking, about the relationship between stress and birefringence in structural proteins of this kind.

There are two types of theories on long-range elasticity, a phenomenological type and a kinetic type. With the same assumptions as to bulk properties (hydrostatic incompressibility and isotropy) the two lines of approach should result in stored-energy functions of the same shape although of different mathematical forms, provided that they represent the same degree of approximation. However, the kinetic theories are of course the more interesting especially as they predict the absolute value of the elastic modulus and because the effect of molecular parameters can be tested in theory and by experiment. The essential feature in the present experiments is therefore not the general similarity between experiment and theory, which may have been spurious, but the demonstration that specified alterations in strain, stress and swelling lead to results predictable from the kinetic theories, including the photoelastic theories. It is therefore necessary to review the main features of these theories.

2. Kinetic Theories of Rubber Elasticity

The kinetic theories have recently been reviewed and discussed in detail by Treloar (1958*a*,*b*) and, with a few exceptions, his notation is used here. Theories deviating only slightly from the ones mentioned in the following have not been considered because physical experiments like the present ones do not offer sufficiently precise numerical answers for refined distinctions.

The kinetic theories are based on three sets of fundamental assumptions:

(1) The unstrained rubber is isotropic, hydrostatically incompressible, and a strain leads to a decrease in entropy but not to any change in internal energy. Physical and thermodynamic experiments show that these assumptions are approximately true for real rubbers, including swollen resilin (Weis-Fogh, 1961).

(2) A rubber consists of an isotropic three-dimensional network of flexible chain molecules which are kept fixed in the network by means of a few stable cross-linkages. The chains between junction points are free and since the individual links can rotate relative to their neighbours (flexibility) each chain is thermally agitated and tends to take up the statistically most probable configuration. The same applies to the assembly of chains, the network. Since the internal energy remains constant a strain only leads to an entropy change, more precisely a change in structural entropy, which can be estimated from the change in distribution of possible configurations which the chains can take up. This molecular model is based partly on assumption (1) and partly on the fact that most rubbers are synthesized from random assemblies of long flexible-chain molecules by some process of cross-linking.

(3) There exists an equivalent random chain for each class of chains. The equivalent chain has the same molecular weight M_c as that of a real chain of the class, the same extended length ($= nl$) and the same root-mean square distance $(\bar{r}^2)^{\frac{1}{2}}$ between the junction points in the randomly kinked condition (Kuhn, 1934). This assumption is a mathematical convenience so as to operate with ideal chains consisting of n links each

of length l, where each link is completely free to point in any direction. In real chains such random links do not occur since the freedom is restricted by valence angles, potential barriers and steric hindrances. An estimate of n, if possible, is therefore a measure of flexibility rather than of the number of actual links and it need not represent a whole number of monomer units.

(a) *Simple kinetic theory; long-chain networks*

If n is sufficiently large and the end-to-end distance r of the equivalent chain is considerably smaller than the extended length nl, i.e. less than about one third, the distribution of end-to-end distances can be expressed by an ordinary Gaussian distribution function with good approximation. The function value is zero for $r = 0$ and has a maximum for $r = l\ (2n/3)^{\frac{1}{2}}$, the most probable length. It is now possible in a simple way to calculate the change in entropy of a single chain when r is altered and, since the change in free energy is determined by the change in entropy, to calculate the tensional force ϕ between the end points of a thermally agitated chain, $\phi = 3kT(r/nl^2)$, where k is Boltzmann's constant and T the absolute temperature. If the expressions are applied to a randomly cross-linked, isotropic network of constant volume, it is possible to express the stored energy as a function of the applied strains and from this the relationship between a unidirectional strain and the tensile force. If f is the tensile force per unit area of the unstrained sample and t is the stress per unit strained area, we have:

$$f = NkT\ (\alpha - \alpha^{-2}) \tag{1}$$

$$t = NkT\ (\alpha^2 - \alpha^{-1}) \tag{2}$$

N is the number of chains per unit volume of the material and α is the extension ratio, i.e. the strained length in units of the unstrained length. Note that n and l do not appear so that N is the only material parameter. In the case of an isotropically swollen material, like resilin, a geometrical parameter v, called the volume fraction, is introduced, where v is the unswollen volume over the swollen volume. It is then convenient also to use the extension ratio $\lambda = \alpha v^{\frac{1}{3}}$ which is the strained swollen length divided by the unstrained swollen length:

$$f_s = NkT\ v^{\frac{1}{3}}\ (\lambda - \lambda^{-2}) = NkT\ (\alpha v^{\frac{1}{3}} - \alpha^{-2}v^{-\frac{1}{3}}) \tag{3}$$

$$t_s = NkT\ v^{\frac{1}{3}}\ (\lambda^2 - \lambda^{-1}) = NkT\ (\alpha^2 v - \alpha^{-1}) \tag{4}$$

The tensile force f_s and the stress t_s now refer to the swollen unstrained and strained areas respectively. The expressions differ from equations (1) and (2) by the factor $v^{\frac{1}{3}}$ (Flory & Rehner, 1943*b*; Treloar, 1947*a*). These equations are of course valid also for networks in which all chains are of equal length (Flory & Rehner, 1943*a*).

In both sets of equations $G = NkT$ is called the elastic modulus. It differs from the modulus of ordinary solids in several respects. It is defined theoretically by the above equations and it is the same for all kinds of deformation at constant volume. It is obvious that it differs from Young's modulus because $(\mathrm{d}f/\mathrm{d}\lambda)_{\lambda=1} = 3v^{\frac{1}{3}}G$ while $(\mathrm{d}f/\mathrm{d}\lambda)_{\lambda\to\infty} = v^{\frac{1}{3}}G$.

The parameter N is related to the number of junction points per unit volume which is $N/2$ in the case of simple cross-linking, where 4 chains meet at each junction point, provided that there are no free ends and no closed loops (network "defects"). It is worth while to recall that, in this context, a junction point may be a stable chemical cross-link as well as a simple physical entanglement. In fact, such entanglements can

make an unvulcanized material consisting of long free polymer chains behave as a rubber at low degrees of stress (*cf.* Moore & Watson, 1956), but it is then subject to flow during prolonged strain and will dissolve completely in a good solvent. (Resilin is completely insoluble.) N is also related to the number-average molecular weight M_c of the chains by the following equation which is valid for simple cross-linking without network defects:

$$G = \frac{\rho RT}{M_c} \tag{5}$$

R is the gas constant (= 84·8 kg cm/degree mole) and ρ the density of the swollen material (for dry resilin 1·3 g cm^{-3}).

The form of the force-extension relationship in equation (1) is seen from Fig. 1, curve A, and since (for large values of α) the slope approaches G it is immediately clear that the

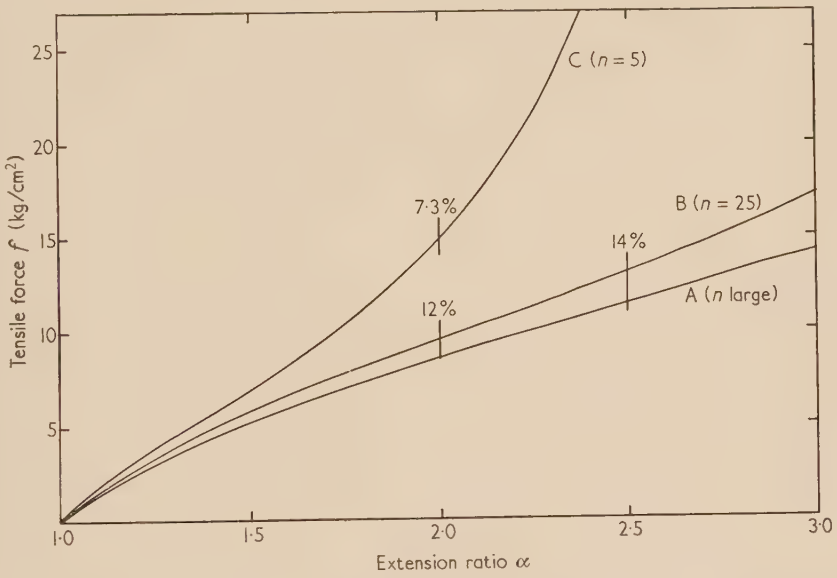


FIG. 1. Theoretical curves showing the tensile force f as a function of the extension ratio α of unswollen isotropic rubbers with (A) a large number n of links in the equivalent random chain (simple theory) and when n is reduced to 25 (B) and 5 (C) respectively. The modulus G is the same in all curves; (B) and (C) are redrawn from Treloar (1954) and refer to the tetrahedral model with non-affine displacement. The increase in force relative to the simple case is indicated for a few extensions.

simple theory fails for large extensions of a network with chains of finite length. However, being a first approximation only, it is slightly inaccurate also at low extensions, offering too high values of f. This is seen *inter alia* by comparison with the Mooney equation which represents the simplest approximation of the phenomenological theory. According to Rivlin & Saunders (1951), we have for a swollen network

$$f_s = 2v^{\frac{1}{3}}(\lambda - \lambda^{-2})\,(C_1 + C_2/\lambda) \tag{6}$$

where C_1 and C_2 are constants. Since we are dealing with swollen resilin where v varies from 0·19 to 0·52, it is important to note (a) that Gee (1946) observed good agreement between experiment and equation (3) in highly swollen rubbers, and (b) that Gumbrell, Mullins & Rivlin (1953) found C_2 in equation (6) to decrease with

swelling and become zero for values of v between 0·2 and 0·3; it was independent of the degree of vulcanization (i.e. of C_1) and of the swelling medium. It should then be possible at small and medium extensions to apply equations (3) to (5) to resilin and from the experiments to estimate the modulus and the approximate number of junction points.

(b) *Short-chain networks*

When the number of links n becomes small and when the extension of the chains increases so that $r > ca.\ nl/3$, non-Gaussian statistics must be applied to calculate the change in entropy of a chain resulting from a change in its length. It is also necessary to specify how the chains are cross-linked and how many random links they contain. Consequently, several theories have been put forward in order to cover the more obvious possibilities. Apart from limited network extensibility they predict that the force-extension curve becomes steeper, as seen from Fig. 1, curves B and C, but the actual shape depends on the specific assumptions made. The tetrahedral model, originally introduced by Flory & Rehner (1943*a*) for long-chain networks, has been used by Treloar (1954) for calculating the entropy change of deformation in a network with few links in the equivalent random chains. It is particularly interesting here since its basic assumptions are much more likely to apply to a protein than to an ordinary rubber which is vulcanized by some random process, for instance by radiation. In the tetrahedral model, all chains of the network are of equal contour length (nl) and of equal properties; the chains are connected so that four chains radiate outwards from the common junction point so that the four neighbouring junction points of this "cell" fall, on the average, at the four corners of a regular tetrahedron. Although Y-branching seems to occur in collagen (Mechanic & Levy, 1959) the common chemical cross-linking in proteins involves 4 half-chains and it is precisely in proteins that a specific and regular distribution of chemical cross-links may be expected due to the long specific amino-acid sequence of the backbone.

Figure 1 shows the force-extension curves, B and C, for a model of this kind, calculated by Treloar (1954) for $n = 25$ and $n = 5$ in a network with non-affine displacement of the junction points of the "cell," i.e. half the total number of junction points are free to take up any possible site within the network. In most network theories the assumption is made that all junction points are moved in correspondence to the macroscopic strain. The non-affine assumption increases the extensibility compared with the affine case but has little effect on the optical properties. In all three curves $G = NkT$ enters as a parameter and has been given the same value of 4·9 kg/cm^2 ($= 6{\cdot}4 \times 0{\cdot}45^{\frac{1}{3}}$, *cf.* p. 657). It is seen that the shapes of the three curves do not differ much at low extensions so that one can obtain an approximate estimate of G by means of the simple theory. But it is also clear that the further course is greatly influenced by the parameter n which, as we have seen, is a measure of chain flexibility.

It should be stressed that this model has not been extended to swollen rubbers (so that the x-axis is in units of α rather than of λ). In principle, the effect of isotropic swelling is to move the starting point further towards the right-hand side so that deviations from the simple theory appear at smaller values of λ than of α. The model has the advantage of having been extended to cover photoelastic effects as well but, again, only a few numerical examples are available.

(c) *Photoelasticity of rubbers*

Most isotropic materials become birefringent when strained but for different reasons. In ordinary solids the strain induces different polarizabilities normal to the stress and in the direction of the stress, due to interatomic displacements. Such changes are correlated with changes in internal energy. In ideal rubbers, the strain-induced birefringence is thought to be due entirely to an orientation of the chain molecules while each link in the equivalent random chain remains unaffected by the stress and has the same optical polarizability α_1 in the direction of its length and normal to it, α_2. For a simple random network of such chains, with many links per chain, Kuhn & Grün (1942) derived a relationship between birefringence $(\mu_1 - \mu_2)$ and strain which was generalized by Treloar (1947*a*) and extended also to include rubbers swollen with an optically neutral solvent (like resilin in water):

$$\mu_1 - \mu_2 = \frac{(\bar{\mu}^2 + 2)^2}{\bar{\mu}} \frac{2\pi N}{45} (\alpha_1 - \alpha_2)\, v^{\frac{1}{3}} (\lambda^2 - \lambda^{-1}) \tag{7}$$

where $\bar{\mu}$ is the mean refractive index of the swollen material, μ_1 the refractive index in the direction of the strain and μ_2 the index normal to the strain. The other symbols are defined as before. If equation (4) is inserted in (7), we get

$$\mu_1 - \mu_2 = t_s \frac{(\bar{\mu}^2 + 2)^2}{\bar{\mu}} \frac{2\pi(\alpha_1 - \alpha_2)}{45\, kT} = t_s C \tag{8}$$

C is the stress-optical coefficient and it is seen that the birefringence is proportional to the stress (force per unit strained swollen area), independent of n and l but inversely proportional to T. It is important that C is also independent of the volume fraction v (apart from v's effect on $\bar{\mu}$). In uncrystallized *Hevea* rubber Treloar (1947*b*) found close agreement between theory and experiments on the first two points and reasonable agreement as to the effect of swelling.

With the same basic assumption with respect to the network as in the simple kinetic theory and with the additional assumption that $(\alpha_1 - \alpha_2)$ is constant and independent of strain, stress, swelling and temperature, equation (8) therefore provides a means of checking the network theories and, conversely, to check whether $(\alpha_1 - \alpha_2)$ remains independent for instance of swelling. The last point is particularly interesting in resilin because it may give valuable information about the stability of the polypeptide chain.

In networks with a reduced number of equivalent links per chain, the stress increases more steeply with strain than in the simple case and the birefringence becomes a non-linear function of the stress, as seen from the unbroken curves in Fig. 10, redrawn from Treloar (1954). They refer to the tetrahedral model with non-affine displacement. In contrast to stress-extension curves, the stress-birefringence curves become less and less steep because the birefringence approaches a finite value (*cf.* horizontal asymptote for $n = 5$). Also, the initial slope decreases with decreasing n and is therefore smaller than that predicted from the simple theory (straight line in Fig. 10).

3. Materials and Methods

The elastic tendons from the metathorax of the large dragonflies *Aeshna cyanea* Müll. and *A. grandis* L. were isolated, stripped of cells and suspended in buffered water in a

small temperature-controlled chamber placed under the polarizing microscope (E. Leitz) as described in a previous article (Weis-Fogh, 1961) in which both the morphology and the stretching technique are discussed.

(a) *Unidirectional compression*

The stripped elastic tendon is cylindrical (0·7 mm long, 0·15 mm in external diameter in neutral saline) and pierced by a central canal (slightly more than 0·03 mm in diameter). By means of fragments of a razor blade the tendon was cut normal to the axis to produce small "drums," 0·2 to 0·3 mm high (Fig. 2 A). Because of imperfect cutting, the distance between the cut surfaces, i.e. the "height" h, varied by a few % over the cross section, so that the entire cross section did not contribute to the force until the piece had been compressed by 5 to 10% of h. The drum was "threaded" by means of a straight piece of platinum-iridium wire, 0·03 mm in diameter (Fig. 2 B). The upper end of the wire was

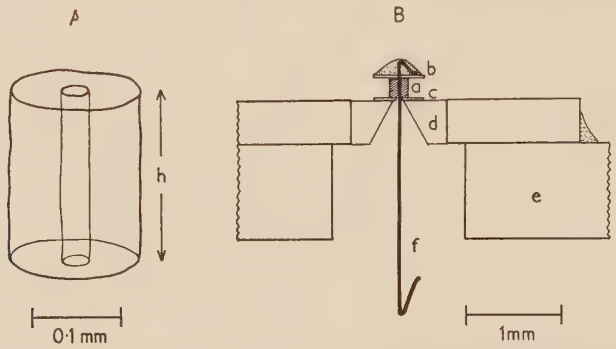


Fig. 2. Method used in compression experiments. (A) Cut resilin "drum" from *Aeshna* tendon, of height h. (B) "Threaded" resilin sample (a) mounted between the upper plate (b) and the sliding plate (c) which rests on the jewel bearing (d). (e) Microscope slide with hole; (f) hook glued to upper plate by means of Araldite (dotted).

fused normal to a plane plate of the same alloy (0·03 mm thick) by means of Araldite. A sliding plate with a hole of the same diameter as the unstrained drum bore was placed beneath the sample. It rested on a pierced, polished surface of a small jewel bearing from a wrist watch which was glued to an ordinary microscope slide. The lower end of the wire was hooked to carry weights placed on the adjustable condensor unit. By lowering the condensor the sample could be compressed by a centered push (up to 4·6 g). The compression was estimated in triplicate by shining light from a small lens-end bulb (without heating) onto minute graphite grains and reading the calibrated fine-adjustment screw after focusing the 45× front lens. The readings were accurate only to within 3 to 4% of h. The sample was repeatedly wetted with buffered water. (In the stretch experiments, the readings of length and force were accurate to within 1%.)

(b) *Birefringence*

The optical retardation Γ was estimated by means of a Berek compensator either with daylight filter (maximum transmission at 590 mμ) or in green monochromatic light at 550 mμ. At all strains both the external and the internal diameters were measured by means of an eye-piece screw-micrometer with accuracy to within 1%. The tendon axis was parallel to the plane of polarization and the retardation was measured normal to the tendon axis so that the optical path touched the central canal tangentially. With this arrangement Γ is a maximum and the errors due to reflection and refraction smallest. The optical thickness d was calculated from the measured diameters in each case (d less than 3% smaller than the external diameter). The birefringence is $\mu_1 - \mu_2 = \Gamma/d$. It is not possible to estimate the accuracy in a simple way but duplicates seldom deviated by more than a few % at the highest retardations (smallest accuracy).

4. Results

As in the thermodynamic experiments (Weis-Fogh, 1961), the tendons reached equilibrium so quickly after the strain that no hysteresis or any other deviation from simple reversibility could be detected either during mechanical or during optical measurements. The strain-induced alterations of the swelling (p. 657) came to an end in one or two minutes but the great variations in v caused by changes in pH were not complete until several minutes after the swelling medium had been altered.

(a) *Unidirectional deformation in neutral water*

Figure 3 shows a typical experiment with a tendon from *Aeshna grandis* (filled circles) at room temperature and pH 6·7. The tendon was first extended in the alphabetical order shown, then cut and subjected to compression. The compression curve is not continuous with the extension curve but is placed too far to the left

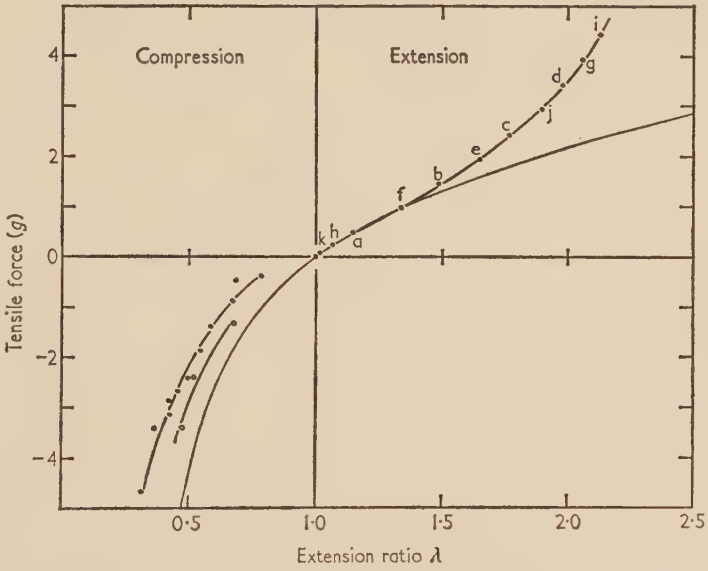


Fig. 3. Typical force-extension experiment with a tendon from *Aeshna grandis*, measured at room temperature and pH 6·7. Ordinate: force (g wt). Abscissa: extension ratio of swollen sample. The alphabetical letters indicate the sequence of measurements. The unbroken curve was calculated according to equation (3).

because the abscissa tends to be wrong by 5 to 10% in that direction, as mentioned on page 654. The few open symbols refer to a tendon from *A. cyanea*. The unbroken curve was drawn according to equation (3); $v = 0{\cdot}38$ was estimated from the swollen dimensions compared with those of the tendon after drying in absolute ethanol, the measurements being made in benzene (Weis-Fogh, 1961); $G = 6{\cdot}9$ kg cm^{-2} was found by making the experimental and the theoretical curves coincide for $\lambda = 0$ and $\lambda = 1{\cdot}2$. It is seen that on the compression side the shape and slope resemble the theoretical ones, but since they are beyond accurate estimation it suffices to stress that resilin is rather similar to ordinary rubber in axial compression. I have not found any data for short-chain networks in compression.

On the extension side, the deviation from the simple course is quite small up to $\lambda = 1{\cdot}4$ but increases progressively at higher extensions. When tendons of similar size were stretched to approximately three times the unstrained length ($\lambda = 3$) they broke, the breaking force being 30 to 40 kg per cm^2 swollen unstrained area. The steep course of the force-extension curve at large strains cannot be explained by changes in internal energy since, even for $\lambda = 2{\cdot}6$, these contributions account for less than 10% of the total force (Weis-Fogh, 1961). It is therefore justifiable to compare the experimental results with those predicted from the short-chain theories, in particular with those based on the tetrahedral model illustrated in Fig. 1.

Three tendons, one from *A. grandis* ($v = 0{\cdot}48$) and two from *A. cyanea* ($v = 0{\cdot}45$), were extended at room temperature (298°K and pH 6·7). The results are seen in Fig. 4 in which the ordinate is the measured force f_s per unit unstrained swollen area

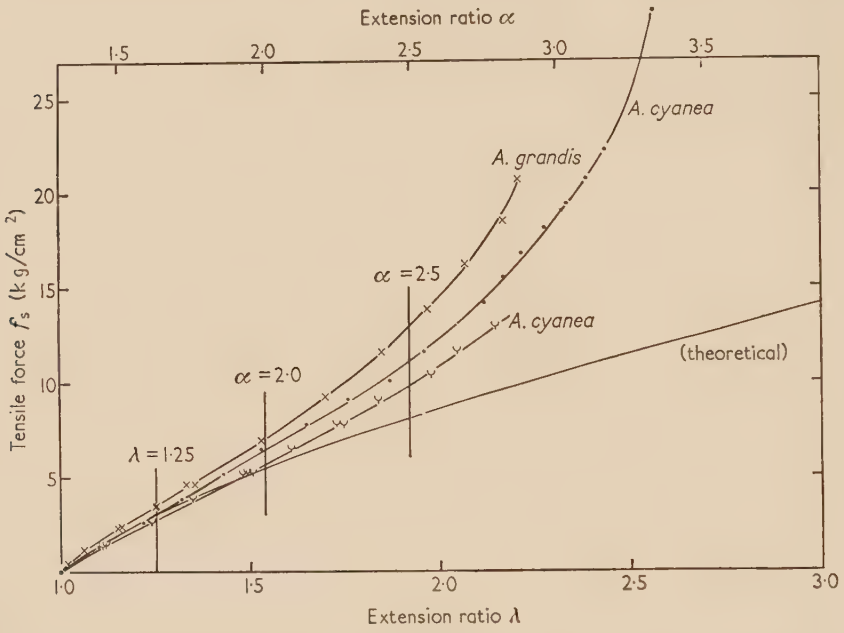


FIG. 4. Force-extension curves of three tendons from two species, measured at room temperature and pH 6·7. Ordinate: force per unit swollen area, f_s. Abscissa: extension ratio based partly on swollen length (λ, lower scale) and partly on dry length (α, upper scale). The modulus G was 7·2 kg/cm^2 for *A. grandis*, 6·4 and 5·7 kg/cm^2 for *A. cyanea*. The unbroken curve was calculated according to equation (3), using the average values for modulus and volume fraction.

and the abscissa is the extension ratio, λ at the bottom (swollen state) and α at the top (unswollen state). Each curve was made to coincide with the theoretical curve in equation (3) for $\lambda = 0$ and $\lambda = 1{\cdot}25$, giving the values of G shown in the legend. The average modulus was 6·4 kg cm^{-2} and the unbroken curve represents the average force-extension relationship for $v = 0{\cdot}46$, as calculated according to the simple theory. The four curves should be compared with those in Fig. 1, although the latter refers to unswollen rubbers. It is then seen that the average force of resilin is increased compared with the simple theory by about 17% for $\alpha = 2$, 28% for $\alpha = 2{\cdot}3$, 37% for $\alpha = 2{\cdot}5$ and by not less than about 80% for $\alpha = 3{\cdot}0$. Corresponding deviations are written against the curves in Fig. 1 in which the modulus was adjusted so that the

"Gaussian" curves are identical in the two figures ($G = 6{\cdot}4 \times 0{\cdot}45^{\frac{1}{3}}$). The deviations from the simple curve in the tetrahedral model are of a character similar to those found in the experiments. The experimental curves correspond to values of n larger than 5 and smaller than 25, an estimate confirmed by the optical analysis.

(b) *The effect of strain and pH on swelling*

In a protein like resilin it is of interest to investigate whether changes in hydrogen-ion concentration influence only the swelling parameter v (simple swelling) or whether they have a more profound effect. It is necessary first to study the effect of the strain itself.

(i) *Effect of strain*

Simple extension of a dry rubber is accompanied by a small decrease in hydrostatic pressure but the effect on the volume is negligible because of the large bulk modulus. In swollen rubbers, on the other hand, the volume may increase appreciably (Flory & Rehner, 1944), roughly as the square root of the extension ratio. If all parameters are taken into account the increase is somewhat smaller. Figure 5 shows that the swelling

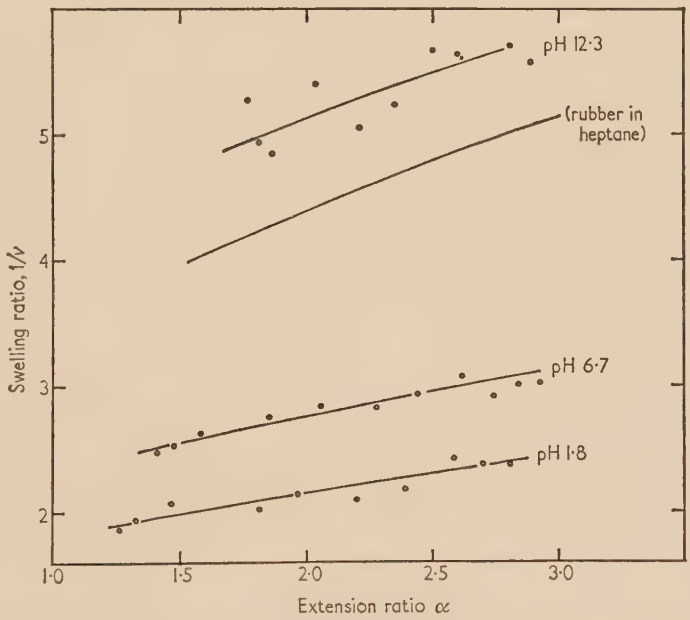


Fig. 5. The effect of simple extension on the swelling ratio $1/v$ of tendon resilin from *A. grandis*, compared with similar experiments with *Hevea* rubber swollen in heptane (redrawn from Treloar, 1950). Room temperature.

ratio $1/v$ of resilin in water at different pH increases with the extension ratio in the same way and to about the same extent as found in vulcanized *Hevea* rubber (swollen in heptane; curve redrawn from Treloar, 1950). In the latter case, Treloar found excellent agreement between experiment and theory, it being assumed that there was no *chemical* affinity between rubber and heptane but only a mutual tendency for diffusion. The experimental values in Fig. 5 were calculated from microscope measurements taken for other purposes so that the accuracy is small but, nevertheless, the

results indicate that the system resilin/water behaves like a simple rubber swollen by an indifferent solvent.

(ii) *Swelling and pH*

This conclusion is corroborated by the experiments summarized in Fig. 6. The stress t_s was measured for each extension and is shown as a function of α for three degrees of initial swelling in a tendon from *A. grandis*. It was placed in each solution for at least 20 minutes before the experiment, first in dilute NaOH, pH 12·3 ($v = 0{\cdot}19$), then in dilute HCl, pH 1·8 ($v = 0{\cdot}58$) and, finally, in 0·01 M-phosphate buffer, pH 6·7 ($v = 0{\cdot}38$). Swelling is minimum at about pH 4 but is nearly as small in the dilute HCl (Weis-Fogh, 1960). According to theory, simple isotropic swelling should affect

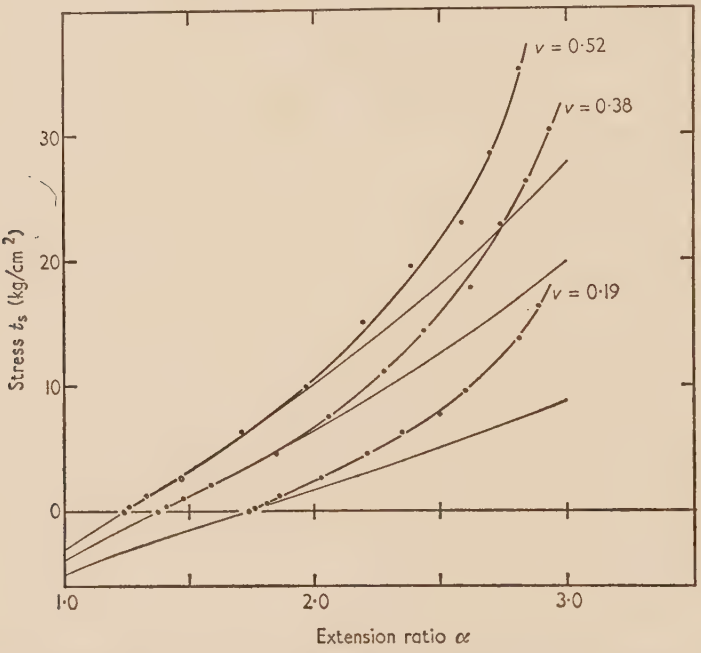


Fig. 6. Effect of swelling on the stress-extension relationship of tendon resilin from *A. grandis*. The volume fraction v was measured in the unstrained state at different pH and the unbroken curves were calculated according to equation (4). Room temperature.

the network only by altering the geometrical parameter v in equation (4). Since the elastic modulus was estimated beforehand to be 6·4 kg cm^{-2}, the real curves could be compared with the theoretical ones, drawn as unbroken lines in Fig. 6. It is seen that the initial parts of the experimental curves closely follow the theoretical curves and that they start to deviate at approximately the same value of α. The drastic changes in pH do not seem to have altered the state of the polypeptide chains in any other way than to have altered their bulk affinity for water, probably by altering the number of ionized sites. This result is supported by the photoelastic experiments.

(c) *Stress birefringence*

The birefringence of resilin, $\Delta\mu = \mu_1 - \mu_2$, is not a linear function of the stress t_s, but the curves are distinctly concave towards the stress-axis, as seen from Fig. 7 in

which two of the curves are displaced along the x-axis by 10 and 20 units respectively. At zero stress the birefringence was always zero. These results are in accordance with Treloar's theory (1954) for networks consisting of short chains of equal length (Fig. 10). In randomly vulcanized dry rubbers, both natural and synthetic, Saunders (1956*a*,*b*) observed a linear relationship even at very high degrees of cross-linking although he found the predicted decrease in initial slope $(\partial \Delta\mu/\partial t)_{\alpha=1}$ for polyethylene and polymethylene but not for *Hevea* rubber.

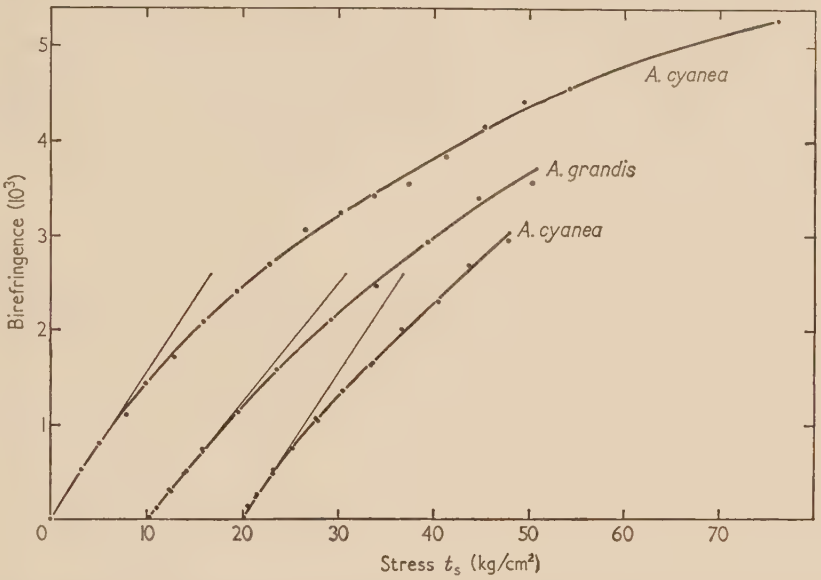


FIG. 7. Stress-birefringence in three tendons from *A. grandis* and *A. cyanea*, measured at room temperature and pH 6·7. Ordinate: birefringence $\Delta\mu \times 10^3$. Abscissa: stress t_s referred to the swollen cross-sectional area. The initial slope indicates the stress-optical ratio C'.

These discrepancies between theory and experiment may be due (a) to the relatively small stresses applied by Saunders, maximally 10 kg cm^{-2} against 76 kg cm^{-2} in my experiments, (b) to the unequal chain lengths which would tend to diminish departures from linearity (qualitative estimate by Treloar, 1954), and (c) to the fact that Saunders used dry rubbers whose stress-strain curves deviate more from the ideal relationship than those of swollen rubbers (Gumbrell *et al.*, 1953). For these and other reasons it is of interest to study the birefringence of resilin in some detail.

(i) *Stress-optical ratio C′*

The stress-optical coefficient C in equation (8) refers to the swollen cross-sectional area and is independent both of the degree of cross-linking and of the volume fraction v while it varies inversely with the absolute temperature. For the time being, it is not possible to apply Treloar's (1954) optical theory rigorously because resilin is rubber-like only in the swollen state for which the theory has not been extended as yet. But according to Saunders (1956*b*, equation (6)), the initial slope in a short-chain dry rubber deviates from C only by a series expansion involving n. If we define the stress-optical ratio of the swollen rubber as $C' = (\partial \Delta\mu/\Delta t_s)_{\lambda=1}$, we may predict that C' should be approximately independent of v, at least within the range of v used here,

and that C' should vary inversely with T. The predictions should only hold provided the applied alterations in pH (swelling) and temperature have no effect either on the number n of links per equivalent random chain or on the polarizability difference of each link $(\alpha_1 - \alpha_2)$, i.e. if the treatments do not alter the state of the polypeptide chains in any other way than by altering the spacing between them in an isotropic manner.

(ii) *Numerical value of C'*

In Fig. 7 the volume fraction for the two tendons of *A. cyanea* was $v = 0{\cdot}45$. At room temperature C' was 1·54 and 1·55 10^{-4} cm^2 kg^{-1}, respectively. It was 1·25 10^{-4} cm^2 kg^{-1} in *A. grandis* ($v = 0{\cdot}48$) but the difference may be accounted for by the larger elastic modulus of this tendon (Fig. 4, upper curve). It is also seen that the curves do not deviate much from linearity up to 10 kg cm^{-2}. These figures seem to be the first estimates of a stress-optical constant in structural proteins.

(iii) *Effect of pH and swelling*

The stress-birefringence was measured in a tendon from *A. grandis* swollen in dilute buffer (from pH 1·8 to 12·3) at room temperature for at least 20 min before measurements started (Fig. 8). The curves are displaced relative to each other by 1 unit

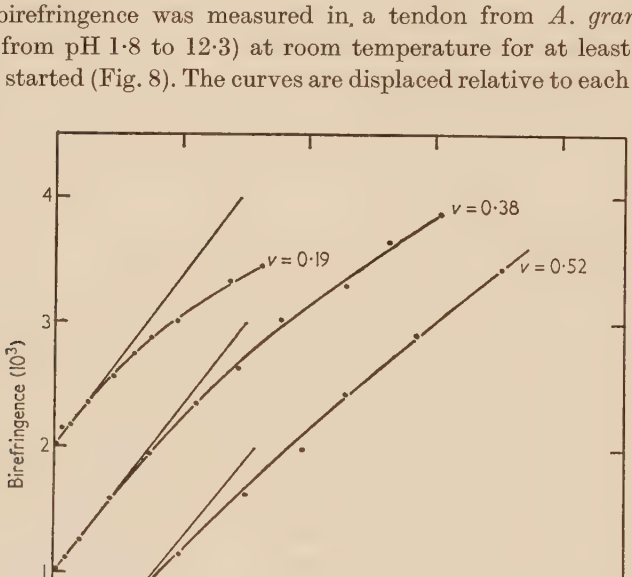


FIG. 8. Three stress-birefringence curves from the same *A. grandis* tendon, swollen to different extents in dilute solutions ranging from pH 1·8 to 12·3. Ordinate: birefringence $\Delta\mu \times 10^3$. Abscissa: stress t_s. Room temperature. C' is almost independent of swelling and of pH.

along the y-axis. C' was 1·26 10^{-4} cm^2 kg^{-1} at pH 1·8 ($v = 0{\cdot}52$), 1·32 at pH 6·7 ($v = 0{\cdot}38$), and 1·36 at pH 12·3 ($v = 0{\cdot}19$). According to equation (8), one would have expected the opposite variation, namely a decrease of about 6%, because the mean refractive index μ decreases with increasing swelling. (In microscope measurements on

dry resilin (Beke lines), $\bar{\mu}$ was 1·56 which for swollen resilin should give $\bar{\mu} = 1{\cdot}333 + v$ 0·24, according to Barer & Joseph (1954).) However, neither the accuracy of measurement nor the theoretical background justify further enquiry. All that can be stated is that C' remained constant to within 10% in spite of the drastic treatment.

(iv) *Effect of temperature*

The stress-birefringence of a tendon was measured twice at room temperature (295°K) and twice at 355°K; the results of one set of estimates are seen in Fig. 9. According to the simple theory, the product $C'T$ should be constant. At the lower

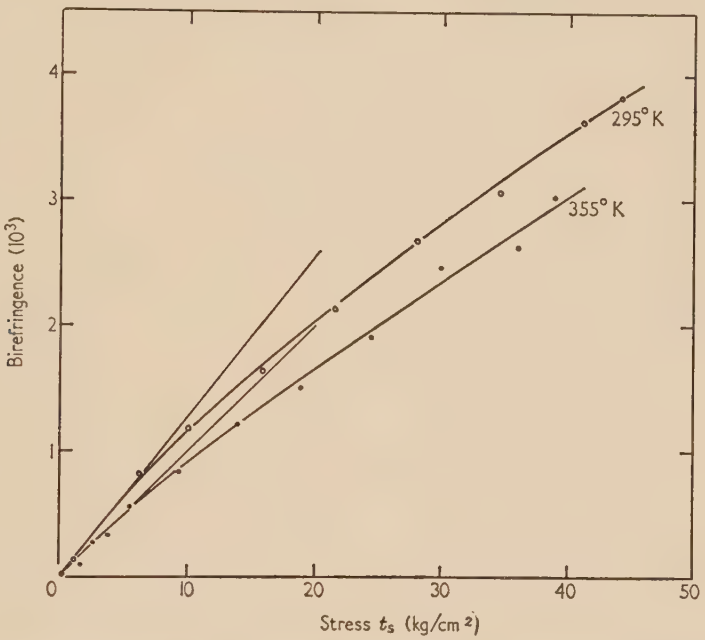


FIG. 9. The effect of temperature on the stress-optical ratio C', as indicated by the initial slope. Tendon from *A. grandis*.

temperature, $C'T \times 10^2$ averaged $(3{\cdot}74 + 3{\cdot}87)/2 = 3{\cdot}81$, and at the higher temperature $(3{\cdot}55 + 3{\cdot}96)/2 = 3{\cdot}76$. The optical properties of the chains did not change, therefore, the effect of temperature on $\bar{\mu}$ being negligible.

(v) *The form of the stress-birefringence curve*

It is now possible with some confidence to compare the experimental results with the photoelastic theory of Treloar (1954) based on the tetrahedral model with non-affine displacement. In his curves, the abscissa is the tensile force $t' = t/NkT$, i.e. the force per unit strained unswollen area divided by the modulus. Consequently, t' was calculated for the upper curve in Fig. 7, using the experimental value of 6·4 kg cm^{-2} for the modulus. In Fig. 10 the straight line corresponds to the simple Gaussian network while the two other unbroken curves refer to $n = 5$ and $n = 25$ in the tetrahedral model. It is seen that the experimental curve coincides with the curve for $= 25$ at the beginning but bends over more rapidly and runs nearly parallel to that for $n = 5$ at high stresses. Bending over of this kind is to be expected if the chains are

of slightly different length because the shortest chains will become dominant at high extensions while the average chain length decides the initial course.

It should be remembered that the theoretical and the experimental curves are not strictly comparable, especially as the stress increases, since resilin was measured in the isotropically swollen state, starting from $\alpha = 1{\cdot}30$ where non-Gaussian phenomena begin to become important. It can be concluded, however, that the behaviour of resilin is consistent with the theoretical model and that the optical data indicate a

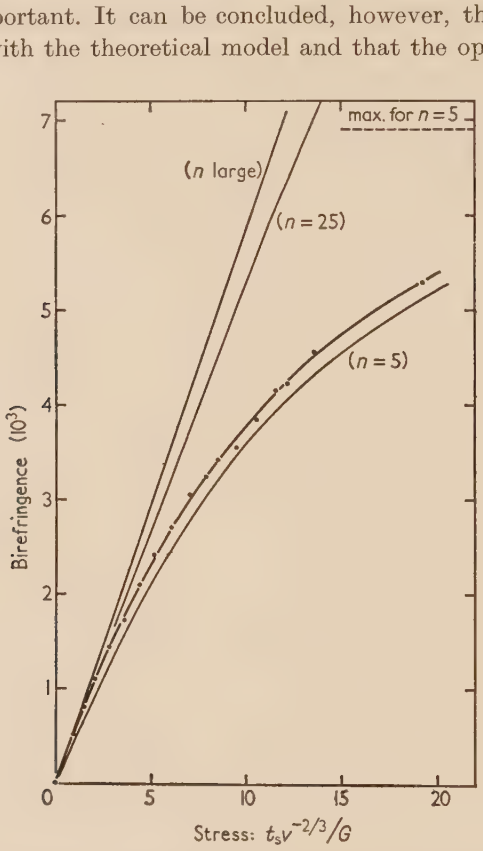


FIG. 10. Comparison between the stress-birefringence curve of a tendon from *A. cyanea* and three theoretical curves (unbroken) calculated by Treloar (1954) for a tetrahedral model with non-affine displacement and with a decreasing number n of links in the equivalent random chain.

flexibility of the backbone corresponding to between 5 and 25 random links in the equivalent chain. The mechanical measurement led to the same result. It should be stressed that none of the results can be due to changes in internal energy or to crystallization which make experiments with other short-chain rubbers so difficult to interpret.

5. Discussion

The mechanical and optical properties of resilin correspond in detail to those of an ideal swollen network consisting of randomly coiled molecular chains of approximately equal length and with relatively few links per equivalent random chain. In fact, no deviation from this model has been observed and the agreement between theory and experiment is even closer than found in natural and synthetic rubbers (Saunders, 1956*a*,*b*), as discussed on page 659. The explanation may well be that resilin and

similar proteins are cross-linked not at random sites but at specific points along the chains. As in other proteins, the specific amino-acid composition of resilin (Bailey & Weis-Fogh, 1961) probably reflects a strict sequence of amino-acid residues and thereby a fixed sequence of actual or potential cross-links, not yet paralleled in other co-polymer rubbers. However, no direct proof of this is to be found in the present experiments. It should be stressed also that network formation based on cross-linking at specific sites rather than on random vulcanization may introduce changes in configurational entropy not encountered in ordinary rubbers (Flory & Rehner, 1943*a*). The probable distribution of chain lengths of the unstrained rubber, swollen or unswollen, may therefore also differ somewhat in the two cases. For these reasons alone the comparison should not be extended too far until direct chemical evidence is available, but it is possible to estimate the approximate number of "active" junction points and to characterize the chain in physical terms.

(a) *The problem of cross-links*

(i) *Evidence for cross-links*

The polypeptide chains of resilin are easily solvated by water and other polar solvents and the swollen protein is rubbery, heat-stable (up to 140°C) and completely insoluble in all solutions which do not break peptide linkages (Weis-Fogh, 1960). The chains must therefore be attached to each other by stable covalent bonds, an interpretation supported by the observed lack of mechanical flow even at tensions approaching the breaking point (Weis-Fogh, 1961). This can either mean that chemical cross-linkages have been formed between the original, unvulcanized protein molecules (inter-chain links) or that they consist of rings which are woven in a three-dimensional network. In the latter case the junction points would all be represented by physical entanglements and there would be no true chemical cross-links but only intra-chain links.

A network made of such rings would satisfy all physical requirements studied here but it is, nevertheless, highly improbable owing to its mode of formation. In order to form an *isotropic* rubber network the chains must either be free in solution and randomly coiled at the time of junction formation or, if the network is formed during the period of linear growth, the greater length of each chain must be free and flexible to slip past the neighbouring chains in a random manner (Flory & Rehner, 1943*a*). In both cases, the probability is extremely small (theoretically zero, *cf.* p. 650) that the free flexible end occupies the same volume element as the other end. It must be stressed that it is necessary for one end to be free in order to dip into another ring before fixation. In fact, to form a three-dimensional network without soluble material (less than 0·3% in resilin; Bailey & Weis-Fogh, 1961) would mean that every chain must get entangled with at least one other chain (linear attachment) and that some chains must entangle two or more chains. Consequently, we must postulate the existence of true interchain cross-linkages of a chemical nature, probably combined with physical entanglements.

(ii) *Number of cross-links*

According to equation (5), an elastic modulus G at room temperature of 6·4 kg cm^{-2} corresponds to a number-average molecular weight M_c of chains between junction points of 5100 g. It is recalled that this estimate relates to networks without defects

and where 4 chains meet at each junction point (p. 650), as in a simple physical entanglement, and also that the estimate is more accurate in swollen than in dry rubbers (p. 651). In an extensive series of experiments with swollen rubbers, Moore & Watson (1956) compared the physical estimates of the number of junction points with the number of chemical cross-linkages known from direct analyses. Over a wide range of cross-linking, the physical determination which relates to the number of effective attachments resulted in 1·5 to 2 times more junction points than chemical cross-links.

In a protein network, one can think of methods of formation which prevent physical entanglements to any appreciable extent. The maximum value for *chemical* cross-linking in *Aeshna* resilin would then correspond to chain lengths with an average molecular weight of $0{\cdot}5 \times 10^4$ g while the more probable value would be 10^4 g.

(b) *Nature of chains between junction points*

The effective chain-molecular weight of tendon resilin is then approximately $0{\cdot}5 \times 10^4$ g. Since the average weight per residue is 88·5 g (Bailey & Weis-Fogh, 1961) there are about 60 residues between two neighbouring junction points. It appears both from the mechanical and from the optical analysis that the probable number of links in the equivalent random chain is greater than 5 and smaller than 25, say about 15. When it is remembered that this corresponds to completely free links with no restriction of freedom, it is clear that the real backbone must contain many more than 15 bonds with rotation. It is generally assumed (*cf.* Pauling & Corey, 1953) that rotation about valence bonds in polypeptides can only take place about the bonds to the carbon atoms which carry the side chains, giving two possible rotational sites per residue (limited flexibility). In the paraffin-type chain with tetrahedral valence angle, for example, each equivalent link (free flexibility) corresponds to not less than 3 carbon-carbon links with free rotation (*cf.* Treloar, 1958*a*). The relatively large side-chains present in proteins, including resilin, and other steric or potential barriers probably mean that the figure is higher for the polypeptide chain. The number of links with rotation then approaches the number of residues in a chain.

In most proteins considerable lengths of the backbone are kept in a regularly folded condition by means of intra-chain hydrogen bonding which accounts for the secondary structure and makes the sections involved stiff. It is on this point that the resilin chains must differ from those of ordinary structural proteins: the amino acid sequence must be such that intra-chain hydrogen bonding is almost prevented whenever the chains are solvated by water, glycol, glycerol, formamide, or by 70 to 80% ethanol in water. In short, swollen resilin lacks a secondary structure. In this connection it is worth recalling that resilin seems completely unable to crystallize so that the X-ray diagram of a maximally stretched and slowly dried tendon is indistinguishable from that of a non-crystalline glass; in the electron microscope it is also devoid of any structure (Elliott, Huxley & Weis-Fogh, 1961). Resilin represents an extreme type of structural protein, the opposite extreme being exemplified for instance by tropomyosin which has a very stable helical configuration, is crystalline and has a distinct banded appearance in the electron microscope (*cf.* Doty, 1959).

Elastin and resilin differ much as to bulk composition, the first having very few, the latter many, polar side groups, but they must be similar with respect to chain flexibility and therefore probably to some fundamental type of amino acid sequence which excludes intra-chain hydrogen bonding.

(c) *General discussion*

(i) *Structure*

One must assume that resilin has a specific primary structure. Since the chains are randomly coiled, since this is the only state observed so far and since it is completely stable, it is concluded that the resilin chains are devoid of any secondary structure. However, the native protein possesses a tertiary structure which extends in *three* dimensions in the form of junction points between chains. Some of these must be chemical cross-linkages and must be so numerous that every chain becomes fixed to the network by means of stable covalent bonds. A piece of resilin is therefore comparable to the space networks known from certain synthetic resins and, in that sense, comparable to a deformable giant molecule with no potential limitations as to size or shape.

For the time being, it is only possible to guess how such a network is formed by living organisms. According to end-group determinations on locust resilin, preliminary estimates by S. O. Andersen (unpublished) indicate that the molecular weight before cross-linking is of the order of 10^5 g. This means that there should be at least 8 to 16 reactive sites along the length of each chain so that it becomes understandable that the final network contains so little soluble material and has such excellent mechanical properties.

(ii) *Extensibility and length*

The number n of links per equivalent random chain determines the extensibility of an ideal network (Treloar, 1954), and, according to this criterion, n for resilin should amount to between 5 and 25, but the estimate is less reliable than the ones based on other criteria because it is very sensitive to the actual distribution of chain lengths. It is, nevertheless, interesting to note that the maximum extension of swollen resilin is similar to that of other rubbery proteins like elastin (Wöhlisch, 1939) and elastoidin II (Fauré-Fremiet, 1936), namely $\lambda_{max} \simeq 3$. The tensile strength is about 30 kg/cm^2 unstrained swollen area in all three proteins, i.e. 70 kg/cm^2 based on the dry area of resilin. Although dry *Hevea* rubber breaks between 200 and 300 kg/cm^2, the strength of the protein is similar to that of vulcanized poly*iso*butylene and silicone rubbers (Roff, 1956). It should also be remembered that swelling and plasticizers tend to decrease considerably the strength in all kinds of rubber so that it can be concluded that there is no decisive difference between rubber-like protein and other rubbers in this respect.

As to the elastic modulus, the lack of information about the volume fraction of swollen elastin and elastoidin prevents detailed comparisons but G is of the same order of magnitude in the three proteins (*cf.* Wöhlisch, 1939; Fauré-Fremiet, 1936). It should be pointed out that unstrained elastoidin II is anisotropic and elastin slightly so (Schmidt, 1937).

(iii) *Biological significance*

There is a general tendency in cytology to think of the solid constituents of cells as being only membranes and filaments made from molecules which are attached to each other by means of more or less stable secondary forces. According to this view the building elements are arranged in one or two dimensions, the three-dimensional cell and its organelles being the outcome mainly of interphase reactions. However, the

space network of resilin may have a more general significance since similar systems may be present as part of living cells, for instance in the cortical areas, in sarcosomes, in muscle fibrils, etc. Such a network, although absent in the major part of the cell, could endow the structures with mechanical stability during stresses of infinite duration whether the stress is of external origin, due to re-arrangements or due to increase or decrease of other cell constituents. A protein network may also provide barriers of the same nature as those used in modern gel filtration systems. Anyway, investigations on the fine structure of cells do not exclude that networks of this kind are present since resilin shows little contrast and no structure in the electron microscope. It tends to escape notice (Elliott *et al.*, 1961).

Since the stress-optical ratio of resilin is practically independent of the degree of swelling and since stresses in soft biological materials seldom exceed 10 kg/cm^2, it may be possible to study in a simple way (*cf.* Figs. 7 and 8) the presence and function of similar protein networks in spite of the difficulty with which they can be identified by other means in living cells.

The experiments indicate two sources of uncertainty which are not recognized in thermodynamic experiments with muscle (*cf.* Hill & Howarth, 1959; Woledge, 1961). First, some proteins like resilin swell more when the temperature is increased, even at constant pH, while the volume of elastin and some other structural proteins decreases (Weis-Fogh, 1961). Second, at constant temperature and pH, the swelling of a protein may change appreciably with strain (Fig. 5) so that the adiabatic change in temperature due to deformation is the outcome of two independent effects, the thermoelastic effect and the concentration effect. It is therefore not possible to estimate the thermoelastic component or even its sign, unless the system is known in detail.

This work was supported by the Department of Scientific and Industrial Research and by The Rockefeller Foundation. I wish to thank Sir James Gray, F.R.S. and Professor V. B. Wigglesworth, F.R.S. for hospitality and research facilities.

Note added in proof

New X-ray evidence obtained by G. F. Elliott indicates that, under special conditions, some crystallinity can be induced in highly stretched, *dry* resilin.

REFERENCES

Bailey, K. & Weis-Fogh, T. (1961). *Biochim. biophys. Acta*, **48**, 452.
Barer, R. & Joseph, S. (1954). *Quart. J. Micr. Sci.* **95**, 399.
Doty, P. (1959). *Proc. IVth Int. Congr. Biochem., Symposium*, **9**, p. 8.
Elliott, G. F., Huxley, A. F. & Weis-Fogh, T. (1961). In preparation.
Fauré-Fremiet, E. (1936). *J. Chim. phys.* **33**, 681.
Flory, P. J. & Rehner, J. (1943*a*). *J. Chem. Phys.* **11**, 512.
Flory, P. J. & Rehner, J. (1943*b*). *J. Chem. Phys.* **11**, 521.
Flory, P. J. & Rehner, J. (1944). *J. Chem. Phys.* **12**, 412.
Gee, G. (1946). *Trans. Faraday Soc.* **42**, 585.
Gumbrell, S. M., Mullins, L. & Rivlin, R. S. (1953). *Trans. Faraday Soc.* **49**, 1495.
Hill, A. V. & Howarth, J. V. (1959). *Proc. Roy. Soc.* B, **151**, 169.
Kuhn, W. (1934). *Kolloidzschr.* **68**, 2.
Kuhn, W. & Grün, F. (1942). *Kolloidzschr.* **101**, 248.
Mechanic, G. L. & Levy, M. (1959). *J. Amer. Chem. Soc.* **81**, 1889.
Moore, C. G. & Watson, W. F. (1956). *J. Polymer Sci.* **19**, 237.

Pauling, L. & Corey, R. B. (1953). *Proc. Roy. Soc.* B, **141**, 21.
Rivlin, R. S. & Saunders, D. W. (1951). *Phil. Trans.* A, **243**, 251.
Roff, W. J. (1956). *Fibres, Plastics and Rubbers.* London: Butterworth Scientific Publications.
Saunders, D. W. (1956*a*). *Trans. Faraday Soc.* **52**, 1414.
Saunders, D. W. (1956*b*). *Trans. Faraday Soc.* **52**, 1425.
Schmidt, W. J. (1937). *Die Doppelbrechung von Karyoplasma, Zytoplasma und Metaplasma.* Berlin: Gebrüder Borntraeger.
Treloar, L. R. G. (1947*a*). *Trans. Faraday Soc.* **43**, 277.
Treloar, L. R. G. (1947*b*). *Trans. Faraday Soc.* **43**, 284.
Treloar, L. R. G. (1950). *Trans. Faraday Soc.* **46**, 783.
Treloar, L. R. G. (1954). *Trans. Faraday Soc.* **50**, 881.
Treloar, L. R. G. (1958*a*). *The Physics of Rubber Elasticity*, 2nd ed. Oxford: The Clarendon Press.
Treloar, L. R. G. (1958*b*). *The Rheology of Elastomers*, 1-14, ed. by P. Mason & N. Wookey. London & New York: Pergamon Press.
Weis-Fogh, T. (1960). *J. Exp. Biol.* **37**, 889.
Weis-Fogh, T. (1961). *J. Mol. Biol.* **3**, 520.
Woledge, R. C. (1961). *J. Physiol.* **155**, 187.
Wöhlisch, E. (1939). *Kolloidzschr.* **89**, 239.

J. Mol. Biol. (1961) **3**, 668–673

The Separation of the Light and Heavy Strands of Bromouracil-substituted Mammalian DNA†

Edward H. L. Chun‡

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.

AND

John W. Littlefield

John Collins Warren Laboratories of the Huntington Memorial Hospital of Harvard University, Massachusetts General Hospital, Boston, Massachusetts, U.S.A.

(*Received 4 March 1961, and in revised form 31 May 1961*)

DNA from mammalian cells which have grown and divided once in the presence of 5-bromodeoxyuridine and in which 5-bromouracil substitutes for approximately one half of the DNA-thymine has an effective density between 1·751 and 1·755 g/cm^3. It can be separated by heat denaturation as shown by density-gradient centrifugation into a light strand which is essentially free of bromouracil and a substituted strand with effective densities between 1·716 and 1·722 and between 1·821 and 1·831 g/cm^3 respectively. These results confirm previous evidence that such DNA is hybrid in nature and that mammalian cells replicate DNA in a semi-conservative manner.

I. Introduction

When various strains of mammalian epithelial cells are incubated either with a relatively high concentration of the thymidine analogue 5-bromodeoxyuridine (BUDR), such as 50 μg/ml., or with lower concentrations in the presence of aminopterin or 5-fluorodeoxyuridine, they appear to synthesize one new complement of DNA in which 5-bromouracil (BU) completely replaces thymine, divide once, and then cease to grow (Hakala, 1959; Cheong, Rich & Eidinoff, 1960; Simon, 1960, 1961; Littlefield & Gould, 1960; Djordjevic & Szybalski, 1960). At lower concentrations of BUDR less complete replacement of DNA-thymine and further cell growth occur. Under these circumstances it has been possible to demonstrate a stepwise increase in the density of DNA during two or more cell growth cycles (Simon, 1960, 1961; Djordjevic & Szybalski, 1960). This suggests that, as in lower organisms (Meselson & Stahl, 1958; Sueoka, 1960), the mode of DNA replication in mammalian cells is semiconservative

† This investigation was supported by research grants C-4186, C-2146 and C-4670 from the National Cancer Institute of the National Institutes of Health, U.S. Public Health Service. This is publication no. 1050 from the Harvard Cancer Commission.

‡ U.S. Public Health Service Research Fellow of the National Institute of Mental Health.

(Delbrück & Stent, 1957). To exclude more conclusively a fragmentative type of replication, the two strands† of BU-substituted DNA (BU-DNA), separated by heat denaturation, have now been examined by density-gradient centrifugation.

2. Methods

Our methods for culturing ERK-1 cells (previously termed KD cells) have been described (Littlefield & Gould, 1960). This epithelial cell line was regarded as derived from rabbit kidney (Westwood, MacPherson & Titmuss, 1957), but appears more likely to be human in origin (Rothfels, Axelrad, Siminovitch, McCulloch & Parker, 1959; Brand & Syverton, 1960; Westwood, personal communication). To obtain BU-DNA about 2×10^7 cells (in the stationary phase of growth because of crowding and lack of refeeding for 3 or 4 days) were distributed into 10 square 8-oz culture bottles in Eagle's medium (Eagle, Oyama, Levy & Freeman, 1957) containing 100 μg/ml. of BUDR (California Corporation).

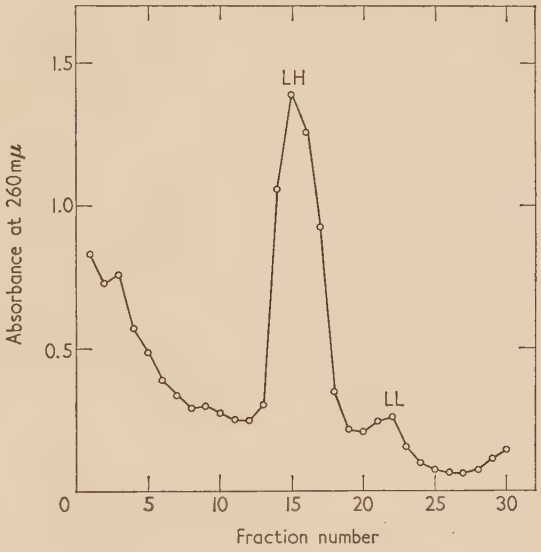


Fig. 1. Absorbance of the diluted fractions of native hybrid BU-DNA XIII collected after preparative ultracentrifugation in CsCl solution. The density decrement per fraction is 0·0083 g/cm³. The density of the major component is 1·755 g/cm³. A small amount of unsubstituted DNA with a density of 1·700 g/cm³ is seen at the right of this peak.

After 3 days incubation, the cells, no longer growing, contained DNA in which 50 to 60% of the thymine was replaced by BU (Littlefield & Gould, 1960). Each monolayer of cells was rinsed with phosphate-buffered saline and lysed by standing for an hour at room temperature in the presence of 0·3 ml. of 1% Duponol (sodium lauryl sulfate). The pooled lysates were dialysed against cold ethylenediaminetetra-acetic acid/phosphate buffer (0·001 M-EDTA, 0·0075 M-Na_2HPO_4, 0·0025 M-NaH_2PO_4, pH 6·8) for 2 days with 2 changes of dialysis fluid and then clarified by centrifugation at 15,000 *g*. DNA was assayed fluorometrically (Kissane & Robins, 1958) or, when isolated, by ultraviolet absorption.

Lysates of samples VIII, XIII, and XIV were further purified by density-gradient centrifugation at 35,000 rev/min for 40 hr at 25°C in an SW39 swinging bucket rotor in the Spinco model L preparative ultracentrifuge. Three ml. of the lysate containing 50 to 100 μg of BU-DNA/ml. and 4·3 g of CsCl were mixed to give a solution density of 1·76

† The term "strand" as used in this paper means the subunits produced by heat denaturation of DNA and does not distinguish whether such subunits consist of one or two chains (Doty, Marmur, Eigner & Schildkraut, 1960; Cavalieri, Rosenberg & Deutsch, 1959).

g/cm³. After the rotor had come to rest without braking, the lusteroid tube was removed from the bucket and punctured at the bottom with a no. 27 G hypodermic needle while under a negative pressure of 2 in. water. The drops, approximately 0·03 ml., were forced out at a convenient rate by gradually increasing the pressure on the tube. Approximately 30 fractions were collected, 5 drops per fraction. Each fraction was diluted with 1 ml. of the EDTA-phosphate buffer and its absorbance at 260 mμ was read in a Cary model 14 spectrophotometer. The BU-DNA was found in 4 or 5 consecutive fractions (Fig. 1) which were pooled and dialysed against the same buffer.

To heat-denature BU-DNA, a sample containing 5 to 50 μg BU-DNA /ml. in the EDTA-phosphate buffer was heated in a water bath at 77°C for 10 min and then quickly cooled in an ice bath. Under these conditions the strands of this DNA should remain separated (Doty, Marmur, Eigner & Schildkraut, 1960). Cesium chloride was added to bring the density to the desired value for density-gradient centrifugation.

The effective densities in CsCl solutions of native BU-DNA and the components of the heat-denatured hybrid BU-DNAs were determined by density-gradient centrifugation in the Spinco model E analytical ultracentrifuge at 44,770 rev/min for 20 hr at 25°C. Tracings of the ultraviolet photographs were made with a Joyce-Loebl microdensitometer. The reference DNA from *E. coli* B was assigned a density of 1·710 g/cm³ (Meselson & Stahl, 1958; Sueoka, Marmur & Doty, 1959).

3. Results

Most of the DNA samples isolated from ERK-1 cells, in which 50 to 60% replacement of thymine by 5-bromouracil could be expected, gave relatively narrow band profiles (Figs. 1 and 2) and had effective densities in CsCl solution in the range 1·751 to 1·755 g/cm³ (Table 1), as compared to a density of 1·700 g/cm³ for DNA from untreated ("normal") ERK-1 cells. Such substituted DNA can be termed "hybrid" (Meselson & Stahl, 1958; see below). Little or no DNA of density of 1·700 g/cm³ was observed in the samples from cells incubated with BUDR, although in a few samples the DNA had broad unsymmetrical band profiles (Meselson, Stahl & Vinograd, 1957).

Table 1

Effective densities in g/cm³ of ERK-1 DNA samples in CsCl solution†

Sample	Native Hybrid	Heat-denatured	
		Light	Heavy
Unsubstituted	1·700	1·716	
	1·700	1·716	
I	1·753	1·716	1·824
			1·826
II	1·753	1·720	
III	1·754	1·722	
IV	1·751	1·716	
VI	1·755	1·719	1·826
VIII	1·753	1·717	1·821
		1·717	1·826
XII	1·754	1·719	
XIII	1·755	1·721	1·830
			1·832
XIV	1·754	1·719	1·825
		1·719	1·826

† Reference DNA samples from *Serratia marcescens* and *Pseudomonas aeruginosa* had densities of 1·7185 and 1·726 g/cm³ respectively referred to the density 1·710 assigned to the DNA from *E. coli* B.

We have not observed a bimodal hybrid band which Davern (1960) found in hybrid BU-DNA of human KB cells. Such a band would result from substitution in DNA in which there were unequal numbers of adenine and thymine in the individual strands.

Ultraviolet absorption photographs of the heat-denatured hybrid BU-DNA at apparent sedimentation equilibrium in a CsCl density-gradient in the analytical ultracentrifuge showed two bands. Figures 2(*a*) and 2(*b*) are densitometer tracings of the "light" (L) and the "heavy" (H) bands along with the reference hybrid (LH)

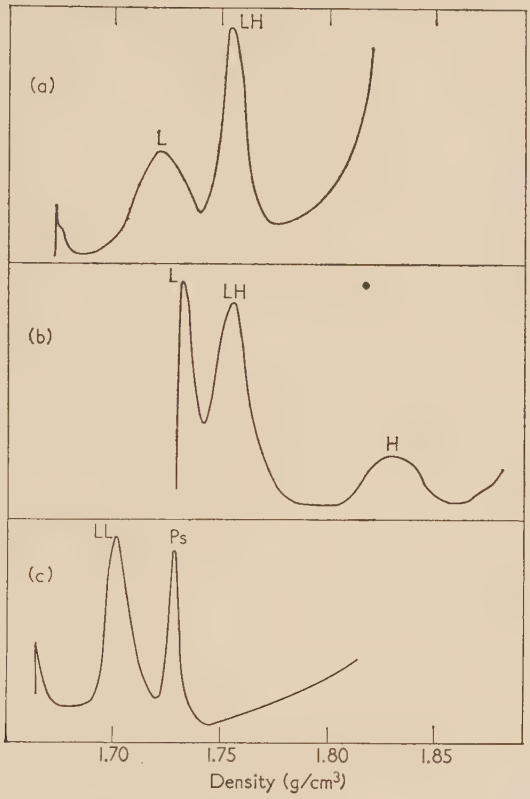


FIG. 2. Densitometer tracings of the ultraviolet absorption photographs of BU-DNA (sample XIII) in the analytical ultracentrifuge cell after 20 hr of density-gradient centrifugation at 44,770 rev/min.

(a) Native hybrid (LH) and the light (L) component of the heat-denatured hybrid.

(b) Native hybrid (LH), light (L) and heavy (H) components of the heat-denatured hybrid. The unresolved light band is at the meniscus.

(c) Native normal ERK-1 DNA (LL) and *Pseudomonas aeruginosa* DNA (Ps).

band. For the various samples shown in Table 1, the effective densities of the light band ranged from 1·716 to 1·722 g/cm³ and those of the heavy band from 1·821 to 1·832 g/cm³. The density of this light band was only slightly greater than that of heat-denatured normal ERK-1 DNA ($\rho = 1{\cdot}716$), indicating very little substitution of BU in the light strand.

Three BU-DNA samples after heat denaturation were centrifuged (35,000 rev/min) to apparent equilibrium (40 hr) in the SW 39 swinging bucket rotor of the preparative ultracentrifuge, and the CsCl gradient thus established was separated into about 30 fractions. From the absorbance at 260 mμ of the diluted fractions two separate peaks

of about equal areas were observed, as shown in Fig. 3, corresponding to the light and heavy bands described above. From refractive index readings of drops of CsCl solution centrifuged under similar conditions, it was established that the density increment was 0·0083 g/cm^3 per fraction. The difference in density measured between maxima of the two peaks, about 0·10 g/cm^3, in the three samples compares favorably with that observed with the analytical ultracentrifuge as shown in Table 1.

To confirm that BU was primarily in the heavy peak, most of lysates I, II, and III were pooled, heated at 85°C for 15 min, cooled rapidly, centrifuged to equilibrium in a CsCl gradient and separated into 14 consecutive fractions as above. Although the absorbance at 260 mμ of the fractions did not indicate in this case a clean separation of the heavy and light strands, several fractions from each peak were pooled (totalling over half the DNA), precipitated and washed with cold 5% trichloroacetic acid, washed with alcohol, incubated at 37°C with N-NaOH, and reprecipitated and washed with cold acid. The base composition of each was determined as previously (Littlefield & Gould, 1960). Corrected to 1·00 for adenine, the composition of DNA from the light peak was adenine 1·00, guanine 0·76, cytosine 0·70, thymine 0·81, and BU 0·13. That for the heavy peak was adenine 1·00, guanine 0·77, cytosine 0·73, thymine 0·14 and BU 0·81.

4. Discussion

The appearance of light and heavy components of the presumed hybrid BU-DNA upon heat denaturation indicates that the hybrid consists of unsubstituted and BU-substituted subunits.

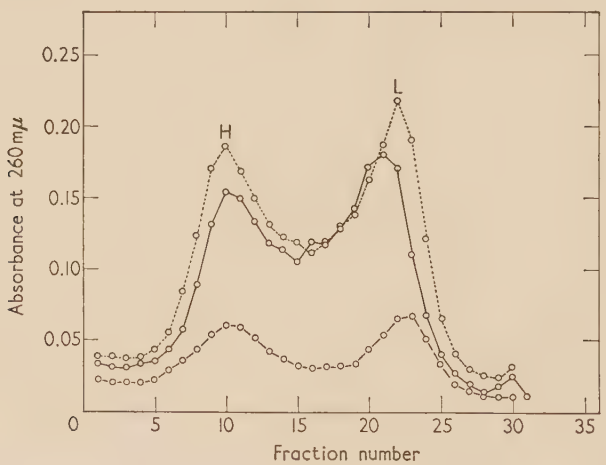


Fig. 3. Absorbance of the diluted fractions of three heat-denatured fast-cooled BU-DNA samples after preparative ultracentrifugation in CsCl solution. – – – –, sample VIII; . . ., sample XIII; ———, sample XIV. Fractions XIII-9 and XIV-9 had densities of 1·830 and 1·831 g/cm^3 respectively.

The essentially equal yield of light and heavy material from the heat-denatured hybrid as seen in Fig. 3 suggests that the hybrid contains equal amounts of each. However, the possibility of unequal recovery of the two components could have led to this result fortuitously. The observed density of the hybrid and of the heavy component released upon heat denaturation furnishes independent evidence for the

presence of equal amounts of light and heavy DNA in the hybrid. Heat denaturation of DNA from various sources results in an increase in density of 0·014 to 0·016 g/cm^3 (Meselson & Stahl, 1958; Sueoka, Marmur & Doty, 1959). In this study heat denaturation produced an increase in the density of 0·016 g/cm^3 in unsubstituted ERK-1 DNA. We assume this value applies also to BU-substituted ERK-1 DNA. Under this assumption, a native molecule containing equal amounts of DNA in its heavy and light strands should have a density of 0·016 g/cm^3 less than the mean of the densities of the heavy and light denatured components. For example, the density predicted for the hybrid of sample I is $\frac{1}{2}(1{\cdot}825 + 1{\cdot}716) - 0{\cdot}016 = 1{\cdot}755$. The observed density is 1·753. The predicted densities agree closely with the values observed for the samples described in Table 1.

We conclude that the hybrid consists of essentially equal amounts of light and heavy strands. This separation of the light and heavy strands of hybrid ERK-1 BU-DNA confirms previous evidence that mammalian cells replicate DNA in a semi-conservative manner.

We should like to thank Professor Alexander Rich for his encouragement and helpful discussions.

REFERENCES

Brand, K. G. & Syverton, J. T. (1960). *J. Nat. Cancer Inst.* **24**, 1007.
Cavalieri, L. F., Rosenberg, B. H. & Deutsch, J. F. (1959). *Biochem. Biophys. Res. Comm.* **1**, 124.
Cheong, L., Rich, M. A. & Eidinoff, M. L. (1960). *J. Biol. Chem.* **235**, 1441.
Davern, C. I. (1960). *Nature*, **188**, 208.
Delbrück, M. & Stent, G. S. (1957). In *The Chemical Basis of Heredity*, ed. by W. D. McElroy & B. Glass, p. 699. Baltimore: Johns Hopkins Press.
Djordjevic, B. & Szybalski, W. (1960). *J. Exp. Med.* **112**, 509.
Doty, P., Marmur, J., Eigner, J. & Schildkraut, C. (1960). *Proc. Nat. Acad. Sci., Wash.* **46**, 461.
Eagle, H., Oyama, V. I., Levy, M. & Freeman, A. E. (1957). *J. Biol. Chem.* **226**, 191.
Hakala, M. T. (1959). *J. Biol. Chem.* **234**, 3072.
Kissane, J. M. & Robins, E. (1958). *J. Biol. Chem.* **223**, 184.
Littlefield, J. W. & Gould, E. A. (1960). *J. Biol. Chem.* **235**, 1129.
Meselson, M. & Stahl, F. W. (1958). *Proc. Nat. Acad. Sci., Wash.* **44**, 671.
Meselson, M., Stahl, F. W. & Vinograd, J. (1957). *Proc. Nat. Acad. Sci., Wash.* **43**, 581.
Rothfels, K. H., Axelrad, A. A., Siminovitch, L., McCulloch, E. A. & Parker, R. C. (1959). *Proc. Third Canadian Cancer Conference*, 189.
Simon, E. H. (1960). Ph.D. thesis, California Institute of Technology, Pasadena, California.
Simon, E. H. (1961). *J. Mol. Biol.* **3**, 101.
Sueoka, N. (1960). *Proc. Nat. Acad. Sci., Wash.* **46**, 83.
Sueoka, N., Marmur, J. & Doty, P. (1959). *Nature*, **183**, 1429.
Westwood, J. C. N., MacPherson, I. A. & Titmuss, D. H. J. (1957). *Brit. J. Exp. Path.* **38**, 138.

J. Mol. Biol. (1961) **3**, 674–683

Degradation of Deoxyribonucleic Acid under Hydrodynamic Shearing Forces

C. Levinthal and P. F. Davison

Biology Department, Massachussetts Institute of Technology, Cambridge, Massachusetts, U.S.A.

(*Received 1 June 1961*)

^{32}P-labeled T2 phage DNA has been subjected to controlled shear forces and the tensile force on the molecule causing scission of the double helix has been determined. Calculation shows that this rupturing force is compatible with the bond strengths. The significance of the molecular weight of DNA normally isolated from organisms is discussed.

1. Introduction

DNA in solution has recently been shown to undergo irreversible degradation when large velocity gradients are produced in the liquid. The sedimentation coefficient and viscosity of the DNA was found to drop after passage through the small orifice of a hypodermic needle (Davison, 1959; 1960) of or a high pressure spray atomizer (Cavalieri & Rosenberg, 1959). An altered chromatographic behaviour and reduced sedimentation rate and viscosity have also been observed after DNA was rapidly stirred in a blendor (Hershey & Burgi, 1960; Rosenberg & Bendich, 1960). It seems likely that these changes represent a drop in molecular weight produced by transverse scission of the elongated polymer. In this paper we will consider the theoretical aspects of shear degradation of such long thin molecules by the action of hydrodynamic shear forces and we will present experiments designed to test some of the quantitative aspects of the theory.

Since our theoretical discussion of this problem assumes laminar flow, at least over the dimensions of the pieces of DNA, it is applicable only to those experiments in which the production of the velocity gradients does not set up turbulence. For this reason the experiments were carried out in a small-bore glass capillary in which flow was maintained at a rate below the critical value necessary to produce turbulence. The ends of the capillary were flared to maintain laminar flow at the entrance and outlet.

The DNA studied was that obtained from the phages T2 and T4 either by osmotic shock or phenol extraction. This material has been shown to have a uniform sedimentation behaviour with a sedimentation coefficient considerably higher than that usually obtained for DNA preparations (Davison, 1959; Hearst & Vinograd, 1961). Furthermore, recent autoradiographic studies have shown that the material obtained by either osmotic shock or phenol extraction has a molecular weight approximately equal to that of the DNA of the intact virus (Davison, Freifelder, Hede & Levinthal,

1961). This large molecule of DNA has a molecular weight which, from published analyses, must lie between 9×10^7 and 13×10^7, corresponding to a contour length of about 45 to 65 μ (see Cummings & Kozloff, 1960).

After these studies were begun Hershey & Burgi (1960) reported that these very long molecules appeared to interact, either directly or through an alteration in the flow pattern, to afford a mutual protection against the shearing forces. It was therefore thought desirable to carry out both the shear and the sedimentation analysis at concentrations of DNA not greater than 0·4 μg/ml. Since DNA cannot be studied at this low concentration in the analytical ultracentrifuge using ultraviolet absorption optics, ^{32}P-labeled phage was used as the source of the DNA and the sedimentation behaviour was studied by zone centrifugation through a stabilizing H_2O-D_2O gradient.

2. Experimental Methods

(a) *Preparation of the DNA*

Unlabeled T2 bacteriophage was grown in Fraser's medium (Fraser & Jerrel, 1953) and purified by filtration through Hyflo (Johns Manville) followed by differential centrifugation. ^{32}P-labeled phage was grown on bacteria which had been incubated in a synthetic medium with $^{32}PO_4$. The phage was purified on a DEAE cellulose column according to the procedures described by Davison *et al.* (1961). This method allows the phage to be purified without being concentrated in a centrifuge pellet and therefore reduces the level of radiation received by the individual phage particles. The DNA was extracted from the phage particles by shaking with phenol (Mandell & Hershey, 1960) or by osmotic shock. The osmotic shock was carried out by placing 0·5 ml. of phage suspension in 4 M-NaCl in a large test tube and rapidly pouring on 9·5 ml. of distilled water.

^{32}P-labeled phage was checked for purity by allowing absorption for 3 min to a suspension of either sensitive or T2-resistant bacterial cells (5×10^8/ml.). In all preparations, more than 95% of the radioactivity remained attached to the sensitive cells when they were separated by low speed centrifugation and less than 5% of the activity sedimented with the resistant cells.

(b) *Shear degradation*

In order to apply high but measurable shear forces to these solutions, they were forced through glass capillaries 7 or 12 cm long, 0·0250 cm internal diameter, by air pressure or by a mechanically driven syringe. As will be discussed in detail below, a number of problems arose from the non-uniform nature of the shear in such a capillary. The more satisfactory method of shearing the solution between rotating co-axial cylinders could not be used since a machine with an annulus small enough to generate the necessary shear without producing turbulence was not available.

(c) *Sedimentation analysis*

(a) For unlabeled solutions containing 10 to 20 μg DNA/ml. the Spinco model E centrifuge equipped with ultraviolet optics was used.

(b) Zone sedimentation was carried out in a Spinco model L centrifuge equipped with a swinging bucket rotor. A linear gradient was formed, using a device similar to that described by Britten & Roberts (1960), between molar solutions of NaCl in H_2O on the top of the tube and D_2O on the bottom (strong salt was added to minimize interaction between DNA and protein). The gradient extended 4·7 cm $\pm$ 0·1 cm up the 5 ml. Lusteroid tube. Above the gradient 0·1 to 0·15 ml. of phage "shockate" in 0·9 M-NaCl was carefully layered. The DNA concentration in the shockate was 0·1 to 0·01 μg/ml. The solution was centrifuged at 35,000 rev/min for 1 hr at approximately 20°C and the contents of the tubes were collected dropwise from holes drilled in the bottom of the tubes.

3. Experimental Results

The efficacy of the zone sedimentation was tested by centrifuging samples of intact phage through stabilizing density gradients. In general, sharp peaks were obtained (Davison *et al.*, 1961, Fig. 1) which served as a check both of the centrifuging procedure and also as an additional test of the purity of the phage preparation. Preliminary experiments also showed that it was possible to resolve sheared and unsheared DNA as shown in Fig. 1. In a few experiments the boundaries were disturbed and the radioactivity was dispersed throughout the tube. Although the reason for this occasional difficulty was not determined, probably thermal gradients disturbed the stabilizing density gradient.

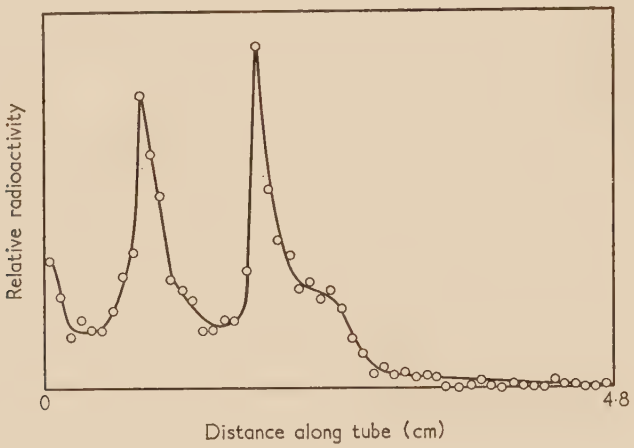


FIG. 1. Distribution of ^{32}P DNA after zone centrifugation of a mixture of sheared and unsheared T2 phage shockate. Direction of sedimentation left to right. The shoulder to the right of the faster peak was present in various proportions in all shockates.

When the DNA solutions were passed through the capillary at a low flow-rate no change was detected in the sedimentation behaviour of the material.

As the flow rate was progressively increased shear degradation was first detected by the appearance of a small second peak of ^{32}P sedimenting more slowly than the intact molecules. At still higher flow-rates a larger fraction of the ^{32}P was in the second peak and this fraction could be increased at any flow-rate by repeated passages through the capillary. These results are easily understandable since the velocity gradient increases linearly with the distance from the axis of the capillary and the maximum shear-force will only be experienced by those molecules which happen to be extended and close to the wall. Thus near the critical flow-rate only those molecules whose centers lie within a narrow annulus near the wall can be degraded. As the flow is increased above the critical rate, molecules within a larger annulus will be broken, but now the molecules in the outer edge of this annulus will experience a force greater than the critical one, and the point of breakage can occur at a distance from the center of the molecule. This distance will increase as the shear forces are raised (see also Burgi & Hershey, 1961). Experimentally it was found that at higher flow-rates fewer passages through the capillary were needed to degrade a given fraction of the DNA.

The intact T2 DNA molecules were found to be degraded at a flow rate of 0·052 ml./sec: no degradation was detected after several passages at 0·043 ml./sec. It may be concluded that the critical flow-rate is close to 0·046 ml./sec.

In order to detect the onset of degradation in the "halved" molecules of DNA, the intact molecules had first to be sheared. For experimental convenience the shockates were sheared at 0·069 ml./sec (50% higher than the critical value). This would be expected to give a greater range in the size of the sheared products than a solution sheared at close to the critical value. However a solution prepared at the higher flow-rate served to indicate the onset of the second critical shear-range when further passage through the capillary produced a new peak containing ^{32}P with a still lower sedimentation velocity. This second critical flow-rate was found to lie between 0·108 and 0·115 ml./sec. Figure 2 shows the sedimentation patterns of a number of sheared samples.

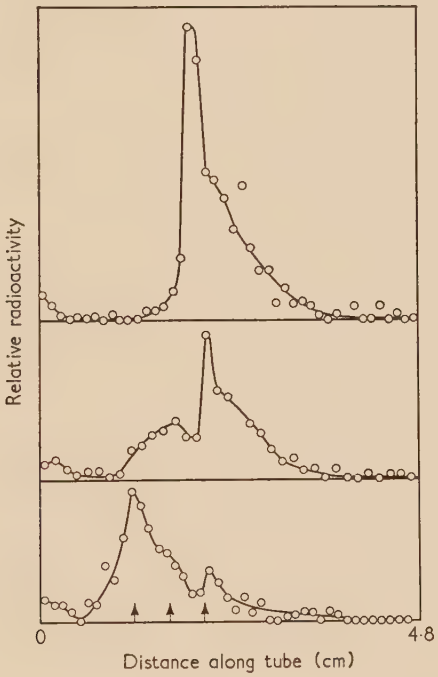


FIG. 2. Distribution of ^{32}P DNA after zone centrifugation of: (upper) T2 shockate sheared at 0·043 ml./sec in 125 μ radius capillary: (middle) similar shockate sheared at 0·09 ml./sec: (lower) similar shockate sheared at 0·12 ml./sec. Direction of sedimentation left to right. Arrows indicate approximate positions of whole, "halved" and "quartered" molecules.

With solutions containing 10 to 20 mg DNA/ml., the initial shear degradation can only be detected at flow-rates between 0·081 and 0·100 ml./sec thus confirming the finding of Hershey & Burgi (1960) that the DNA molecules at concentrations above 4 μg/ml. afford each other mutual protection against degradation by shear forces.

When DNA heavily labeled with ^{32}P was studied, the critical flow-rate fell progressively on successive days to a value below 0·036 ml./sec. It is believed that this was due to the effect of the ^{32}P decays causing the scission of single strands of the DNA molecules, thereby weakening the structure. When DNA labeled at very low specific activity was used this effect was not noticed during the brief periods of these experiments.

A tail was apparent on the fast side of the unsheared DNA sedimentation pattern. This tail varied from preparation to preparation and in one case the diagram appeared bimodal. Since the diagram of any one preparation was consistent it was supposed that the tail was some indication of the presence of phage or contaminating protein or polybases which interacted with or crosslinked the DNA. Some confirmation of this explanation was obtained by centrifuging DNA purified on a column after the methods of Mandell & Hershey (1960). The rapidly sedimenting tail to the distribution was absent in this preparation (Fig. 3).

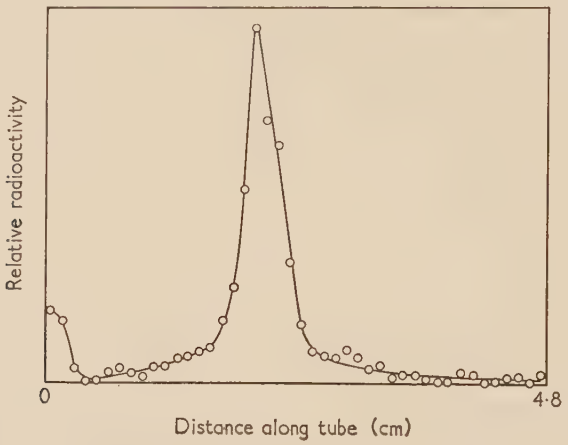


Fig. 3. Distribution of ^{32}P DNA in presence of cold carrier after zone centrifugation. DNA was previously deproteinized by phenol and purified by chromatography.

It was found that, by shearing the solution at approximately 0·057 ml./sec flow-rate three times through a 12 cm capillary, more than 50% of the molecules could be detected in the new slower peak corresponding to the first drop in the sedimentation coefficient. This implies that, if the molecules in the shockate are not monodisperse, more than half of them have roughly the same susceptibility to shearing. Using shearing in a high-speed stirrer and chromatographic fractionation Hershey & Burgi (1960) reached the same conclusion.

4. Theory

The fact that hydrodynamic forces on long polymer molecules could be sufficiently strong to break covalent bonds was first pointed out by Frenkel (1944). He calculated the maximum tension which would be produced in a long chain molecule in a uniform velocity gradient. The model used was one in which the molecule was replaced by a string of spherical beads, each of which was assumed to interact with the fluid independently of the others. However, the frictional drag on an ellipsoid of revolution has been calculated exactly and a good approximation has been carried out for a long cylinder which agrees closely with the value for ellipsoids with a large axial ratio (Burgers, 1938). For these long thin particles the frictional drag is not simply proportional to the length of the molecule as would be the case for non-interacting spheres. Instead there is a logarithmic dependence on the axial ratio, which, for the very long molecules we are considering, makes a significant (more than a factor of 10) difference to the calculated force on the molecule. In the following calculations we will assume

the frictional drag per unit length (f) as constant along the length of the rod but we will use the theoretical value of the total resistance P appropriate for a rod of the length in question.

The frictional resistance to a long rod moving through a viscous medium is (Burgers, 1938)

$$P_1 = 2\pi\eta\, L \Big/ \left(\log\frac{L}{\rho} - 0{\cdot}72\right)$$

and

$$P_2 = 4\pi\eta\, L \Big/ \left(\log\frac{L}{\rho} + 0{\cdot}5\right)$$

where P_1 and P_2 are the forces opposing motion when the rod moves parallel and normal to its axis, respectively, η is the viscosity of the solvent, and L and ρ are the length and radius of the rod. For a rod inclined at 45° to the stream-lines we have arbitrarily chosen a value between these limits. Hence we assume the frictional resistance per unit length as:

$$f = 3\pi\eta \Big/ \left(\log\frac{L}{\rho}\right) \tag{1}$$

A fully extended molecule inclined at an angle θ to the stream-lines in a uniform velocity gradient G experiences a stretching force τ which is maximal in the center of the molecule and equal to:

$$\int_0^{L/2} fGL \sin\theta \cos\theta \, \mathrm{d}L$$

τ is greatest when $\theta = 45°$

$$\tau_{\max} = fGL^2/16$$

Turning to the case of flow through a cylinder of radius a, the velocity profile of an incompressible liquid is:

$$v = \frac{2F}{\pi}\left(\frac{a^2 - x^2}{a^4}\right) \tag{2}$$

where v is the velocity of the liquid at a distance x from the axis of the tube, and F is the flow-rate, volume per unit time. Therefore in a cylinder the velocity gradient is no longer uniform but increases linearly with x. Solute molecules near the cylinder wall will experience the greatest shearing stress, and this force will be maximal at a point slightly displaced from the center of the molecule. To calculate the maximum stress imparted to any molecule at a given flow-rate, consider a fully extended molecule in a plane intersecting the axis of the cylinder and at 45° to the stream-line, with one end in contact with the wall.

Changing the coordinate origin from the axis to the cylinder wall by substituting $a - x = y$ in (2) (Fig. 4)

$$v = \frac{2F}{\pi} \cdot \frac{2ay - y^2}{a^4} \tag{3}$$

Let the molecule move with velocity v_0. The liquid stream moves with velocity v_0 at $y = y_0$.

Let $L/\sqrt{2} = y_{\mathrm{m}}$. Since opposite stretching forces are equated about y_0:

$$\int_0^{y_{\mathrm{m}}} f(v - v_0)\mathrm{d}y = 0$$

or

$$3y_0^2 - 6ay_0 = y_{\mathrm{m}}^2 - 3ay_{\mathrm{m}} \tag{4}$$

In these experiments the radius of the tube was 125 μ but the appropriate length L is uncertain between 45 to 65 μ. In each of the calculations which follow the results will be given for the case of $L = 65$ μ. The point of breakage of a 65 μ long rod can be calculated from (4) to be 0·9 μ from the center of the rod. For a halved molecule ($L = 32$ μ) the point of breakage will be less than 0·2 μ from the center. If the velocity of flow were a function of higher powers of distance, as might be the case in turbulent flow, the point of breakage might be further removed from the center.

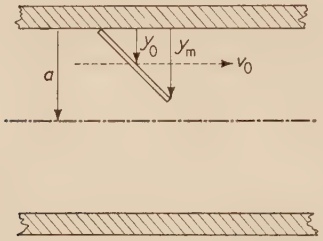


FIG. 4. Model situation considered for quantitative analysis of forces imposed by hydrodynamic shearing.

The maximum stress within the molecule will be

$$\tau_{max} = \int_0^{y_0} f(v - v_0)dy$$

$$= \frac{2fF}{\pi} \cdot \frac{y_0^2}{a^4}\left(a - \frac{2y_0}{3}\right) \tag{5}$$

Experimentally the critical flow-rates which just produced degradation of the whole and "halved" molecules in the 125 μ radius capillary were 0·046 ml./sec^{-1} and 0·11 ml./sec^{-1}. Substituting in (5) we can compare the stresses τ_W and τ_H needed to degrade the whole and half molecules. τ_W/τ_H is found to be 1·35. The fact that this ratio is greater than 1 is not unexpected in view of the fact that the "half" molecules were produced at well above the critical shear and hence some of the molecules would be considerably longer than 33 μ. In fact calculation shows that this ratio is exactly unity if the maximum length of the molecules which could be produced at $F = 0{\cdot}069$ ml./sec is inserted in the equation. However, these details are of minor importance in view of the approximations and assumptions we have made. It is possible that such elongated molecules as DNA are never completely elongated at 45° to the stream-lines, but become extended at lower angles as suggested by the model experiments of Forgacs & Mason (1959).

The magnitude of the tension causing scission of the whole molecules can be calculated from equations (1) and (5) to be $1{\cdot}1 \times 10^{-3}$ dynes. The tension which the DNA backbone may be expected to withstand may be calculated by differentiation of the potential function relating to the two atoms involved in the most vulnerable bond. Lippincott & Schroeder (1955) proposed an approximation to the potential function which has proved useful for polyatomic molecules:

$$U(r) = D\left(1 - \exp\left(\frac{-n(r - r_0)^2}{2r}\right)\right)$$

where n is a factor determined by the atoms joined by the bond, D is the dissociation energy and r_0 is the equilibrium bond length. The force exerted by the bond is dU/dr.

$$\frac{dU}{dr} = \frac{Dn}{2}\left(1 - \frac{r_0^2}{r^2}\right)\exp\left(\frac{-n(r - r_0)^2}{2r}\right) \tag{6}$$

This force is maximal where $d^2U/dr^2 = 0$. This requires:

$$\frac{nr_0}{4}\cos^4\theta = \sin^3\theta \text{ (where } \sin\theta = r_0/r)$$

The rupture of any of three types of bonds, C—C, C—O or P—O could cause the scission of the DNA chain. We need only consider the weakest of these bonds, and hence we can exclude the P—O bond which probably has some double-bond character. Using equation (6) and the values of n, D and r_0 appropriate to the C—O and C—C bonds (Pauling, 1960) it was calculated that these bonds could resist a tensile force of $8{\cdot}9\times10^{-4}$ and $8{\cdot}1\times10^{-4}$ dynes respectively.

5. Discussion and Conclusions

The stress on a molecule in the velocity gradient of a capillary can be calculated with reasonable precision, if our assumption is correct that during the course of a few passages there is a high probability that the molecule will become fully extended and at 45° to the flow lines. *A priori* this assumption is not unreasonable since the particles will tumble of the order of 200 times while passing through our capillary at the flow-rate required to give the critical shear, and it is only necessary for one of these tumbling motions to result in full extension. However, more direct support of this assumption is obtained from the experimental results. At the critical shear the scission is produced near the center of the molecule as indicated by the change in the sedimentation coefficient as well as by the fact that the sedimentation pattern is still rather sharp. Even when the shearing is done at a rate of 50% above the critical value there is little overlap between the sedimentation pattern of the degraded and undegraded material. This result is incompatible with the view that the molecule is still in a random configuration at the instant of breakage. Furthermore, the existence of a critical shear for breakage seems to require that the molecule pass through some unique configuration at which it suffers maximum stress. The ratio of the flow-rate required to produce the first break to that required to break the first degradation product is found experimentally to be in reasonable agreement with that calculated assuming in both cases that the particles are fully extended.

With the assumption that the molecule is a single unbranched filament which is fully extended when breakage occurs the calculation of the shear force depends only on the validity of the hydrodynamic resistance for a long cylinder. The various approximate methods used by Burgers for this problem give good agreement with the exact calculations for the ellipsoid of revolution, both for transverse motion of the particles and for rotation about the center of the particle. Thus we conclude that the calculations of the magnitude of the stress is reasonably accurate as long as turbulence is avoided. The critical flow-rate necessary to produce turbulence in a pipe is reached when the Reynolds number† is between 1600 and 1700. Below this value any turbulence produced will be damped out, while above this value turbulence can spread through the pipe. In our case the value was maintained below 1200.

The calculations of the maximum stress which a covalent bond can withstand is uncertain, due to lack of information about the exact parameters of the bonds involved as well as uncertainties in the exact shape of the potential to be used for the particular bonds involved. On the other hand none of the uncertainties can be

† Reynold's number for a pipe is the mean fluid velocity × the pipe diameter × the liquid density/viscosity of fluid.

very large, since the calculation is relatively insensitive to the shape of the potential and depends primarily on the bond distance and maximum energy of the bonds. Therefore within a factor of 2 the maximum tension can reasonably be taken as 9×10^{-4} dynes. The fact that after some ^{32}P decays have taken place in the molecule it is more susceptible to shear indicates that both chains of the double helix must be broken at the same time. Thus the total force required to break the molecule would be between 9 and 18×10^{-4} dynes, which is in good agreement with the value of 11×10^{-4} dynes, for the stress produced by the flow gradient at the critical shear.

The experimental results presented here as well as those of others indicate that DNA may be degraded in the course of purification and manipulation and that the previously measured molecular weights of DNA do not necessarily have any relationship to the molecular weights of the material prior to extraction. Neither the uniformity of the material obtained nor the consistency of the values found in any one laboratory indicate that the material is unchanged by the extraction. Carefully controlled procedures will generally lead to approximately the same shear forces being applied to the molecules, which will in turn produce uniform and consistent molecular weights. If trout sperm DNA is extracted in such a way as to minimize shear (Davison, 1960) the observed sedimentation coefficient is very much higher than the values previously reported. This indicates that the shear degradation process is not limited to the DNA of the T2 virus. The fact that the maximum tension required to break the molecules is in fair agreement with the estimates of the covalent bond strength, indicates that the backbone of the DNA molecule itself is probably being broken by the shear.

In the case of the T2 phage all the DNA can be shown to be in one piece after extraction and the breakage can be accounted for as scission of internal bonds in the DNA. Therefore it is clearly not necessary to postulate non-nucleic acid linkers in the structure of the chromosome. However, turning to the DNA of larger chromosomes, the question immediately arises as to whether there are any considerations which set an upper limit to the amount of DNA which can exist as a single molecule. The estimation of the DNA content of the nuclei of human cells gives as the total value about 4×10^{-12} g. Since this material is contained in 46 chromosomes, the total length of the DNA in a single chromosome would be a few centimeters. Can one muster any arguments (other than intuitive incredulity) against the possibility that all of this material forms a single DNA molecule? One might assume that the double-stranded nature of DNA with the requirement for unwinding on division, could immediately eliminate an enormous structure as a single unit duplicating as an entity. If we assume along with Longuet-Higgins & Zimm (1960) that the unwinding process is unrelated to duplication then the time required to unwind a molecule of length 2 cm, due to the change of configurational entropy, would be impossibly long even if we accept their assumption that the strength of the hydrogen bonds can be reduced to zero. On the other hand, it is possible that the unwinding and duplication are related processes such that the bulk of the energy released in cleaving the nucleotide triphosphates is available for the mechanical work required to rotate the molecule in a viscous medium. With this assumption the calculations of Levinthal & Crane (1956) indicate that even a molecule of length 2 cm could be made to rotate sufficiently rapidly so that the DNA duplication could be accomplished in about 10 hours. This is approximately the time observed for DNA duplication in mammalian cells and it is during this interphase period that the visible chromosome structure does not exist in the cell. In making these calculations we assume that about 8 kcal are available

per nucleotide inserted in the new DNA and that this energy is used to counteract the frictional drag involved in rotating the long cylinder in a medium with a few centipoise viscosity. This result is not altered if the unwinding ends are condensed into a compact ball. Since the viscous drag varies as the square of the radius, the energy required for rotation is the same for a very long thin rod as for an equal mass of material condensed into a shorter thickened structure.

Certainly, none of these considerations demonstrates that single molecules of length 2 cm or more actually exist in cells. However, they do point out that the intuitive conclusions based on our usual experience with macroscopic mechanics are not applicable in considering problems of intracellular motion. Since the lability of a very long molecule to shear increases extremely rapidly with its length, the hypothesis that the DNA of a chromosome is all contained in one molecule seems almost impossible to confirm by direct observation. However, we can conclude that the molecular weights of DNA in such cells are probably very much greater than the usually quoted value of 6 to 10 million molecular weight. In the case of phage DNA of molecular weight about 120 million the physical properties are those one could expect from a single structure and there is no indication of any weak links in the backbone (Hershey & Burgi, 1960) which might have been expected if breaks occurred in the DNA chain, or if protein segments were inserted into the structure.

Part of the work was performed in the laboratory of Professor Francis O. Schmitt for whose helpful co-operation the authors are grateful.

This research was supported by grants E-2028 and E-1469 from the National Institute of Allergy and Infectious Diseases.

REFERENCES

Britten, R. J. & Roberts, R. B. (1960). *Science*, **131**, 32.
Burgers, J. M. (1938). *Kon. Ned. Akad. Wet., Verhand.* **16**, 113.
Burgi, E. & Hershey, A. D. (1961). *J. Mol. Biol.* **3**, 488.
Cavalieri, L. F. & Rosenberg, B. H. (1959). *J. Amer. Chem. Soc.* **81**, 5136.
Cummings, D. J. & Kozloff, L. M. (1960). *Biochim. biophys. Acta*, **44**, 445.
Davison, P. F. (1959). *Proc. Nat. Acad. Sci., Wash.* **45**, 1560.
Davison, P. F. (1960). *Nature*, **185**, 918.
Davison, P. F., Freifelder, D., Hede, R. & Levinthal, C. (1961). *Proc. Nat. Acad. Sci., Wash.* in the press.
Forgacs, O. L. & Mason, S. A. (1959). *J. Colloid Sci.* **14**, 473.
Fraser, D. & Jerrel, E. A. (1953). *J. Biol. Chem.* **205**, 291.
Frenkel, J. (1944). *Acta Physicochim., U.R.S.S.* **14**, 51.
Hearst, J. E. & Vinograd, J. (1961). *Arch. Biochem. Biophys.* **92**, 206.
Hershey, A. D. & Burgi, E. (1960). *J. Mol. Biol.* **2**, 143.
Levinthal, C. & Crane, H. R. (1956). *Proc. Nat. Acad. Sci., Wash.* **42**, 436.
Lippincott, E. R. & Schroeder, R. (1955). *J. Chem. Phys.* **23**, 1131.
Longuet-Higgins, H. C. & Zimm, B. H. (1960). *J. Mol. Biol.* **2**, 1.
Mandell, J. D. & Hershey, A. D. (1960). *Analytical Biochem.* **1**, 66.
Pauling, L. (1960). *The Nature of the Chemical Bond.* Ithaca, N.Y.: Cornell University Press.
Rosenberg, H. S. & Bendich, H. (1960). *J. Amer. Chem. Soc.* **82**, 3198.

J. Mol. Biol. (1961) **3**, 684–692

The Nature of the "Deletion" Mutants in the *r*II Region of Phage T4

Masayasu Nomura† and Seymour Benzer

Department of Biological Sciences, Purdue University, Lafayette, Indiana, U.S.A.

(*Received 29 June 1961*)

Certain *r*II mutants of phage T4 behave as if a sizable segment of the genetic structure has been affected. Experiments were done to determine whether this segment is deleted or is still present in some altered form. The results of genetic crosses, showing a shrinkage of the distance between outside markers, indicate that the deletion hypothesis is correct.

1. Introduction

Some *r*II mutants of phage T4 (Benzer, 1955) behave as if they contain mutations that extend over segments of the genetic map. That is, they give no detectable recombinants when crossed with any of several other mutants which do give recombinants with one another. This type is found most frequently among mutants that do not revert detectably. Their properties can be simply understood as owing to the deletion of a segment of hereditary material corresponding to the map span covered, and for this reason such mutants have been loosely referred to as "deletions", in analogy (on an intra-cistronic scale) to the chromosomal deficiencies that are so well known in higher organisms (Bridges, 1917). However, other models would do as well, for instance (a) a series of many mutations or replacement of the normal structure by a stretch of "nonsense," or (b) an inversion of the normal segment, since, as in higher organisms, an inversion would be expected to lead to a greatly diminished frequency of recombination when the mutant is crossed with others having mutations within the affected region. Models (a) and (b) differ from a deletion in that genetic material corresponding to the altered segment would still be present in the mutant. The present study was undertaken to determine whether or not this is the case for extended *r*II mutants.

2. Materials and Methods

Phage T4B was used. *r*II mutants designated as *r* without any other prefix were derived directly from T4B by spontaneous mutation. *r*H23 and *r*H88 are spontaneous mutants derived from a revertant of a point *r*II mutant. *r*NB3093 and *r*NB3365 (courtesy of Dr. I. Tessman) were induced with nitrous acid.

Strain B of *E. coli* was used as the host for crosses in the genetic experiments, and for plating. Strain K (*E. coli* K12(λ), Benzer's strain) was used as the selective plating

† Present address: Institute for Protein Research, Osaka University, Osaka, Japan.

indicator for detection of standard type recombinants. The broth medium was 1% Difco bacto-tryptone plus 0·5% NaCl, with agar added for plating (1·2% for the bottom layer, 0·7% for the top layer).

The procedures for preparing phage stocks and for standard crosses were those described by Chase & Doermann (1958). The multiplicity of infection was usually between 10 and 15 of each parent. The recombination value, when K was used as indicator, is expressed as twice the ratio of plaques on K to plaques on B, corrected for the efficiency of plating of r^+ on the two strains.

Spot tests for crossing *r*II mutants were performed as follows: one drop of each mutant stock at a titer of about 10^9/ml. was added to 0·5 ml. of a culture of strain B (about 10^8 cells/ml.). After allowing a few minutes for adsorption, a drop of the mixture was spotted on an agar plate which had been seeded with K, and the results were scored after overnight incubation. A negative result (no plaques on the K background) in this test indicates that less than about one in 10^5 recombinant progeny are produced in the cross.

CsCl density-gradient centrifugation (Meselson, Stahl & Vinograd, 1957) was performed in an attempt to detect a difference in density between standard type phage and extended mutants. Since the density of T4 particles varies depending on the conditions of growth (Champe, unpublished experiments), each mutant was grown in mixed infection with standard type under one step growth conditions using *E. coli* strain BB, which does not discriminate between the standard and mutant types. The lysate was mixed with CsCl solution in broth to give a final phage titer of about 2×10^6/ml. and a density of 1·50. Three ml. of this solution was heated at 45°C for 2 hr, a treatment which improves the sharpness of banding (D. Pratt, personal communication). After cooling, the sample was centrifuged in a Spinco SW39 swinging bucket rotor at 25,000 rev/min for 24 hr. The plastic tube was withdrawn and pierced at the bottom with a needle. Each drop was collected in a separate tube and suitably diluted to assay both r^+ and r type phages, using strain B as indicator.

3. Results

(a) *Cross between an extended mutant and standard type*

Figure 1 shows the relative locations in the genetic map of the various mutants used in this study, as determined by recombination experiments. The mutant *r*1272 does not produce standard type recombinants when crossed with any known *r*II mutant. Thus it is drawn as entirely covering both cistrons of the *r*II region.

If the mutation in *r*1272 consists of a series of many mutations, a stretch of nonsense, or an inversion, it should be possible, by crossing it to the standard type, to obtain recombinants containing only a portion of the alteration. To test for such splitting, *r*1272 was crossed with standard type and the progeny plated on strain B. Approximately half the progeny were *r* type and, of these, 740 plaques were picked and a stock made from each. These were crossed by spot tests against *r*205 and *r*287, which contain well-separated mutations in different cistrons. All but one failed to give any recombination with either *r*205 or *r*287. One showed recombination with both, and was found by further crosses to be a point mutant; it presumably arose by mutation from the standard type, rather than recombination. Therefore, less than 0·2% of the progeny of the cross showed any splitting of the *r*1272 mutation. This result is in agreement with expectation for a deletion but is in contrast to the expectation of over 4% (based on the frequency of recombination between *r*205 and *r*287) if *r*1272 were a series of many mutations or a string of nonsense. In the latter case, however, the frequency of recombination might be reduced by the difficulty of pairing between standard type and mutant in the *r*II region, so that this result is not conclusive.

(b) *Examination of heterozygotes involving an extended* r*II mutation*

In crosses between T4r^+ and r mutants a small proportion of the progeny particles are heterozygous, i.e. carry both r^+ and r markers which segregate in the progeny. This heterozygosity applies to only a small part of the phage genome, so that the probability that a particle will be heterozygous for two mutations decreases as the distance between the mutations increases (Hershey & Chase, 1951; Levinthal, 1954).

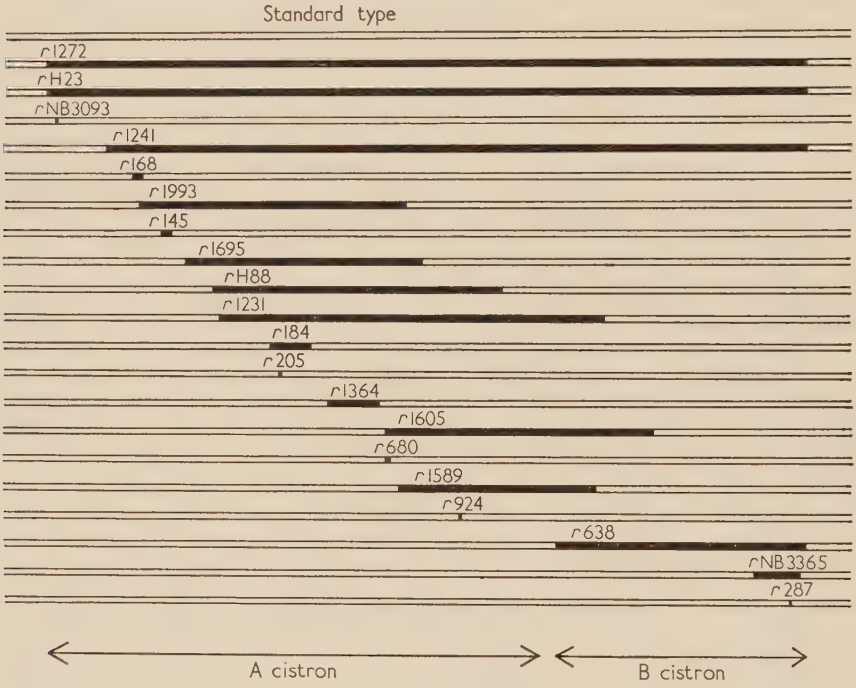


FIG. 1. Genetic map of the rII region of phage T4 showing the relative positions and extensions of the mutants cited in this paper.

Figure 2 illustrates the heterozygote types to be expected from a cross between standard type and a double mutant having two well-separated rII markers. Most of them will be types (a) and (b), i.e. heterozygous for only one of the two markers. Therefore, splitting of a genetic segment may be much more frequent among heterozygotes than among the general progeny of a cross. If an extended mutant, covering both of these markers, is not a deletion, but contains an altered structure of some sort, the heterozygotes it produces in a cross with standard type could include some containing only a portion of the original lesion. Such heterozygotes could be detected by crossing their r-type segregants with two distant point mutants within the segment; if an r segregant from a heterozygote has lost a part of the original extended mutation that overlaps one of these point mutants, it will give recombination with that mutant. On the other hand, if the extended mutation is a deletion, the r type segregants of heterozygotes should always contain the entire original mutation.

r1272 (which covers the entire rII region) was crossed with standard type and the progeny were plated on strain B. Mottled plaques were picked, replated and stocks

made from six r type segregants from each mottled plaque. These were examined for their genotype by testing against two rII mutants ($r205$ and $r287$) which are separated by 4·4% recombination units (Chase & Doermann, 1958). As a control, the same experiment was done using the double mutant $r205$ $r287$ in place of $r1272$.

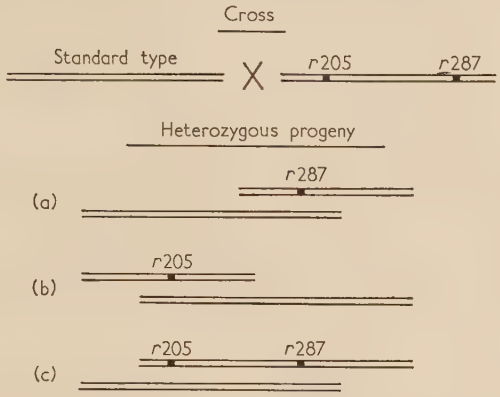


FIG. 2. Various possible structures of heterozygotes arising in the cross of a double rII mutant with standard type. All of these will give rise to mottled plaques.

The results are shown in Table 1. In the control experiment, out of 69 mottled plaques examined, 65 were found to be heterozygous for only one of the mutations, either $r205$ or $r287$, and 4 were heterozygous for both. On the other hand, in the cross between $r1272$ and r^+, out of 80 mottled plaques none gave recombination with either $r205$ or $r287$. (Four exceptional plaques were found, but these were shown to be due

TABLE 1

Heterozygotes produced in crosses of standard type with (A) a double r*II mutant and (B) an extended mutant*

Source of heterozygotes	Cross A $r^+ \times r205\ r287$	Cross B $r^+ \times r1272$
Total mottled plaques examined	69	80
Number of heterozygotes yielding rII segregants that recombine with $r205$ (but not with $r287$)	36	0
Number of heterozygotes yielding rII segregants that recombine with $r287$ (but not with $r205$)	29	0
Number of heterozygotes yielding rII segregants that recombine with neither $r205$ nor $r287$	4	76
Others	0	4†

† These contain neither $r205$ nor $r287$, but other r mutations altogether (in two cases not even rII mutations). They presumably derive from mutations on the plate during the development of r^+ plaques. Standard type phage alone produced about 0·04% mottled plaques compared with 0·53% in the cross $r^+ \times r1272$ and 1·1% in the cross $r^+ \times r205\ r287$. Therefore, about two such spurious mottled plaques are to be expected in Cross A and about six in Cross B.

to newly arising mutations.) Thus, no splitting of the extended mutation was observed, a result consistent with the deletion hypothesis.

A fact worthy of mention is that extended mutants showed a reduced frequency of heterozygote production compared with point mutants. (See Appendix.)

(c) *Shrinkage of distance between outside markers*

As can be seen in Fig. 3, the recombination values between two outside markers and the extended mutation *r*1695 add up to much less than when they are crossed with each other. This confirms the notion that a sizeable part of the structure in *r*1695 is altered, but does not, of course, distinguish between any of the proposed models.

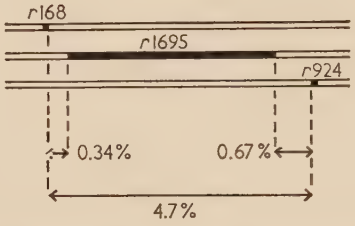


FIG. 3. Recombination frequencies observed in crosses involving the entended mutant *r*1695 and two mutants having nearby outside markers.

This can be done by means of double mutants as shown in Fig. 4. The mutants are made homologous for the *r*1695 mutation, so that, regardless of the nature of that mutation, recombination within the segment should be normal. If *r*1695 is a deletion the distance between the two outside points should be correspondingly shorter. On the other hand, if *r*1695 consists of a multiply-mutated, inverted, or "nonsense" segment no such shrinkage should be observed.

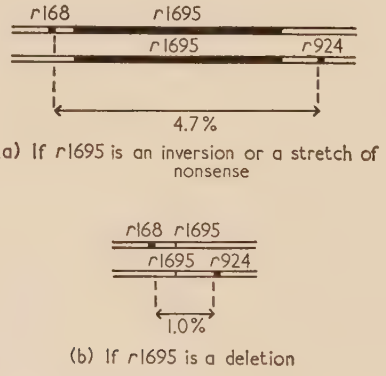


FIG. 4. Experiment to determine whether *r*1695 is a deletion. The two double mutants (obtained by recombination) are homologous for the *r*1695 mutation. Thus, if *r*1695 is not a deletion the distance between outside markers should be unchanged. The recombinants to be measured are *r*1695 and *r*168 *r*1695 *r*924.

The double mutants *r*168 *r*1695 and *r*1695 *r*924 were isolated from the progeny of the appropriate crosses by testing *r*-type progeny against the parental types. To make certain that the entire *r*1695 mutation was incorporated the double mutants

were tested against three mutants (*r*184, *r*1364 and *r*680) spaced within the segment covered by *r*1695.

The double mutants were crossed and the frequency of recombination between the two outside markers, i.e. the probability of getting either *r*1695 or the reciprocal recombinant *r*168 *r*1695 *r*924, was measured. Since both recombinant types are *r*-type, K cannot be used for their selection. Each progeny plaque was picked, grown into a stock and spot tested against *r*168 and *r*924. The proportion of recombinants anticipated among the progeny, if *r*1695 is a deletion, is, from Fig. 4,

$$0{\cdot}34 + 0{\cdot}67 = 1{\cdot}0\%$$

On the other hand, the alternative models predict that the frequency should remain the same as in the cross between *r*168 and *r*924, i.e. 4·7%.

The results of the analysis of about 2,000 progeny phages are described in Table 2. The frequency of recombination observed between the outside markers was 0·7%, close to the value predicted by the deletion hypothesis.

TABLE 2

Effect of the r*1695 mutations upon distance between the outside markers* r*168 and* r*924*

Cross A *r*168 × *r*924		Cross B *r*168 *r*1695 × *r*1695 *r*924	
Total progeny analysed	1104	Total progeny analysed	2007
Parental types		*Parental types*	
*r*168	549	*r*168 *r*1695	1004
*r*924	497	*r*1695 *r*924	960
Recombinant types		*Recombinant types*	
*r*168 *r*924	20	*r*168 *r*1695 *r*924	6
r^+	†	*r*1695	8§
Total recombinants	40‡	*Total recombinants*	14
Percent recombination between outside markers = 3·6 ± 0·8		*Percent recombination between outside markers with* r*1695 present* = 0·7 ± 0·2%	

Shrinkage of distance due to *r*1695 = 2·9 ± 0·8%

† 38 progeny gave positive reactions with both *r*168 and *r*924 when crossed by spot tests. These included both true r^+ and various kinds of heterozygotes, but further analysis to distinguish between these was not performed.

‡ This value was obtained by doubling the number of *r*168 *r*924, on the assumption that the reciprocal recombinant is equally frequent.

§ In addition there were 29 which behaved like *r*1695 in the spot test (i.e. gave recombination with both *r*168 and *r*924) but were shown by further testing to be heterozygotes. Among these, 23 segregated both parental types, while 6 segregated *r*1695 and one parental type.

As an appropriate comparison, the frequency of recombination in a cross between *r*168 and *r*924 was measured using the same technique, that is scoring the double mutant recombinants by a spot test. The value obtained was 3·6% (Table 2).

This is somewhat smaller than the value of 4·7% obtained by doubling the proportions of progeny producing plaques on K, but some difference is to be expected due to the contribution of heterozygotes. The recombinant heterozygotes of the structures *r*168$^+$ *r*924$^+$/*r*168 *r*924$^+$ and *r*168$^+$ *r*924$^+$/*r*168$^+$ *r*924 can grow on K and give rise to an *r*$^+$ plaque, while the corresponding reciprocal classes of recombinant heterozygotes *r*168 *r*924/*r*168 *r*924$^+$ and *r*168 *r*924/*r*168$^+$ *r*924, cannot (Edgar, 1958). The latter types would give positive results against either *r*168 or *r*924 in the spot test, and therefore would not be scored as double mutants. The same consideration may also apply to the small discrepancy between the observed value 0·7 and the value 1·0 predicted from the deletion hypothesis in the cross between *r*168 *r*1695 and *r*1695 *r*924. Six heterozygotes which segregate the *r*168$^+$ *r*1695 *r*924$^+$ type recombinant were in fact scored in this cross (Table 2). Using these data, the sum of recombinants and recombinant heterozygotes can be calculated to be $0{\cdot}7 \times \frac{8+6}{8} = 1{\cdot}2\%$.

(d) *Comparison of the density of mutant particle with standard type*

The analysis described above indicates that some *r*II mutants lack a segment of the genetic map. If this corresponds to a deleted segment of hereditary material composed of DNA and the size of segment is an appreciable portion of the total, the altered DNA content of such mutants might be reflected in a reduced density of the mutant phage particles. This has been successfully demonstrated by Weigle, Meselson & Paigen (1959) for defective mutants of phage λ, using the CsCl density-gradient centrifugation technique of Meselson *et al.* (1957).

Several *r*II mutants of T4 were examined, as described in Methods: *r*1272, *r*H23, *r*1241, *r*145, *r*1993, *r*H88, *r*1231, *r*1695, *r*1605, *r*1589, and *r*638. In each case, *r*II mutant phage banded at exactly the same position as standard type. The density difference between standard type phage and the various *r*II mutants tested was in every case less than that corresponding to successive drops, or less than 0·001 g/ml.

4. Discussion

While the genetic data confirm the deletion hypothesis, the density gradient experiment failed to show a density difference between standard type and extended mutants of 0·001 g/ml. or greater. From this figure, the maximum possible decrease in DNA content (assuming the amount of protein to remain constant) can be calculated, following Weigle *et al.* (1959). Taking the measured value 1·5 g cm^{-3} for the buoyant density of T4 and the value of 1·7 g cm^{-3} for T4 DNA (Meselson *et al.* 1957), the decrease in DNA content is calculated to be less than 1%. This implies that the *r*II region constitutes no more than 1% of the DNA. Based on the determination by Hershey, Dixon & Chase (1953) of the total DNA per particle, this would place an upper estimate of 2,000 nucleotide pairs on the size of the *r*II region. Other possibilities cannot be excluded, however. Deletion of a segment of DNA in the *r*II region might lead to concomitant loss of an equivalent amount of protein, or the deleted piece might be translocated to some other part of the genetic structure. In these cases no density change of the particle would be observed.

A minimum estimate for the size of the *r*II region can be made from the fact that it is known to contain over 400 distinct parts separable by recombination (Benzer,

1961) and presumably consists of at least that number of nucleotide pairs. This is five times below the limit of sensitivity of the density gradient method as used here, so that there is not necessarily any contradiction between the failure to observe density differences and the conclusion that the mutations are deletions.

A more sensitive method for determining the physical size of a deletion, analogous to the classical experiments of Bridges, Skoog & Li (1936) involving deletions at the notch locus in *Drosophila*, may be direct observation, in the electron microscope, of a DNA duplex in which one strand comes from standard type while the other comes from a deletion mutant. Such duplexes might be prepared by the method of Marmur & Lane (1960) for separation and reassociation of the DNA strands. In this case, the segment in the standard type strand corresponding to the deletion should appear as an unpaired loop.

McFall & Stent (1958) described a "star" mutant of phage T2 found among survivors of ^{32}P decay. It gave a low frequency of recombination with two markers which are ordinarily distantly linked. Therefore, the star mutation seems to extend over a considerable sector of the T2 genome. However, the nature of that mutant is clearly different from the ones described here. In crosses of the star mutant with wild type they were able to replace various sectors of the mutation with their normal counterparts, so that the alteration of the genetic structure in the star mutant is very likely in the nature of an inversion. It would therefore appear that both deletions and inversions, well known in the chromosomes of higher organisms, can also occur in phage.

We wish to call attention to the fact that Dr W. Bode (1962) has performed experiments similar to ours with the extended mutants *r*H88, *r*1589 and *r*1231. In each case a shrinkage of the distance between outside markers was observed. It seems likely, therefore, that most *r*II "deletions" are, in fact, deletions.

This research was supported by grants from the National Science Foundation and the National Institutes of Health.

REFERENCES

Benzer, S. (1955). *Proc. Nat. Acad. Sci., Wash.* **41**, 344.
Benzer, S. (1961). *Proc. Nat. Acad. Sci., Wash.* **47**, 403.
Bode, W. (1962). Doctoral dissertation, University of Cologne, Germany (to be published).
Bridges, C. B. (1917). *Genetics*, **2**, 445.
Bridges, C. B., Skoog, E. N. & Li, J.-C. (1936). *Genetics*, **21**, 788.
Chase, M. & Doermann, A. H. (1958), *Genetics*, **43**, 332.
Edgar, R. (1958). *Genetics*, **43**, 235.
Hershey, A. D. & Chase, M. (1951). *Cold Spr. Harb. Symp. Quant. Biol.* **16**, 471.
Hershey, A. D., Dixon, J. & Chase, M. (1953). *J. Gen. Physiol.* **36**, 777.
Levinthal, C. (1954). *Genetics*, **39**, 169.
Marmur, J. & Lane, D. (1960). *Proc. Nat. Acad. Sci., Wash.* **46**, 453.
McFall, E. & Stent, G. S. (1958). *J. Gen. Microbiol.* **18**, 346.
Meselson, M., Stahl, F. W. & Vinograd, J. (1957). *Proc. Nat. Acad. Sci., Wash.* **43**, 581.
Weigle, J., Meselson, M. & Paigen, K. (1959). *J. Mol. Biol.* **1**, 379.

APPENDIX

Depression of Heterozygote Formation by Extended Mutations

When a point *r*II mutant of T4 is crossed with standard type, approximately 1·4% of the progeny produce mottled plaques due to heterozygosity for the *r* marker. Edgar (personal communication) discovered that the frequency of heterozygotes for one extended mutant (*r*164) was lower than for a point mutant. This led us to study other mutants to see whether the effect is general.

Extended *r*II mutants of various lengths were crossed with standard type and the progeny were examined for the frequency of mottled plaques. The results, which confirm Edgar's observation on *r*164, are shown in the table.

*Proportion of heterozygotes among the progeny from crosses of various r*II *mutants with standard type*

*r*II mutant	Relative length of the mutation	Total plaques examined	Number of † heterozygotes	Percentage of heterozygotes
*r*131	1	8,510	119	1·40 ± 0·13
*r*NB3093	1	3,103	44	1·42 ± 0·20
*r*164	20	11,520	52	0·45 ± 0·06
*r*NB3365	22	3,774	18	0·48 ± 0·12
*r*1695	93	10,784	47	0·44 ± 0·06
*r*1272	308	29,706	158	0·53 ± 0·04
Control: Standard type alone		7,675	3	0·04 ± 0·02

Each *r*II mutant was crossed with standard type, using a multiplicity of infection of 14 particles per cell of each parent. The relative length of a mutation is expressed in terms of the number of known genetic sites overlapped. *r*131 and *r*NB3093 are, so far as is known, point mutations.

† Each mottled plaque was picked and replated to confirm its content of both *r* and r^+.

The smallest extended mutant tested (*r*164) already shows the full reduction in frequency. This suggests that there may exist two classes of heterozygotes, one class which cannot form with an extended mismatch and a second class which is insensitive to mismatch.

The effect was also apparent in the case of a double mutant having one extended and one point mutation. When tested separately, *r*NB3093, a point mutant, yielded about 1·4% heterozygotes in a cross with standard type, while *r*NB3365, an extended mutant, gave a heterozygote frequency of about 0·5%. When the double mutant *r*NB3093 *r*NB3365 was crossed with standard type, it yielded about 1·1% mottled plaques. Sixty one such mottled plaques were examined to determine the structure of the heterozygotes; 49 showed heterozygosity for the *r*NB3093 mutation only, 4 for the *r*NB3365 mutation only, and 8 for both mutations. This result, obtained within a single cross, confirms, at least qualitatively, the difference observed between point and extended mutants.

J. Mol. Biol. (1961) **3**, 693–698

LETTERS TO THE EDITOR

The Effect of 8-Azaguanine on the Ribonucleic Acid of the Ribosomes in *Bacillus cereus*

In a previous paper (Otaka, 1960), it was reported that 8-azaguanine, when added to a growing culture of *Bacillus cereus*, almost completely stops the formation of both the protein of the ribosomes and the soluble protein. Under the same conditions, the increase in RNA of the ribosome fraction is reduced to about 30% of the control. The synthesized RNA which contains azaguanine is accumulated in the ribosome fraction, thus resulting in an increase in the RNA/protein ratio of this fraction.

In the present paper, we report experiments indicating that this increase in azaguanine-containing RNA in the ribosome fraction reflects an accumulation of precursor RNA of the ribosomes. A small amount of the precursor RNA is also detected in the normal ribosomes. Another point of interest will be that, in the growing bacterial cells, the "DNA" peak as separated by column chromatography always contains some RNA, presumably in the form of "DNA-RNA hybrid." This RNA is accumulated when the ribosomal RNA synthesis is slowed down by the addition of azaguanine.

An exponentially growing culture of *Bacillus cereus*, NCTC 569, in casamino acid-salt medium at 37°C was divided into three parts, which received (1) [8-^{14}C]-adenine (0·02 μc = 0·76 μg/ml. culture) for 60 min (Exp. A, control), (2) first [^{12}C]-azaguanine (40 μg/ml. culture) for 30 min and then [8-^{14}C]-adenine (0·02 μc = 0·76 μg/ml. culture) for 60 min in the presence of [^{12}C]-azaguanine (Exp. B), or (3) [2-^{14}C]-azaguanine 0·07 μc = 40 μg/ml. culture) for 90 min (Exp. C). The cells were washed, ground as previously described (Otaka, 1960), and extracted with 3 vol 5 mM-tris buffer (2 amino-2 hydroxymethylpropane-1: 3-diol), pH 7·4 containing 5 mM-Mg acetate. The 12,000 *g* supernatant was centrifuged at 105,000 *g* for 120 min to separate the ribosomes (Rb) and the supernatant fraction (Sp). Resuspended Rb, or the Sp, was then fractionated on a diethylaminoethyl-cellulose (DEAE) column (1·2 cm × 15 cm, exchange capacity: 0·44 milliequivalents/g) with a linear concentration gradient of NaCl (*cf.* Bolton, Britten, Cowie & Roberts, 1958).

Figure 1 illustrates a typical elution pattern of the control Rb (Exp. A). Ribosomal particles (P-I peak), which represented about 85% of the u.v. absorbing material, were eluted at salt concentrations between 0·3 and 0·6 M. The remaining 15%, which contained RNA but no DNA, was eluted with alkali (P-II peak). The proportion of the alkali fraction to the ribosomes greatly increased in the azaguanine-containing culture, as seen in Figs. 2 and 3. In Exp. A, in which nearly all the available [^{14}C]-adenine in the medium had been exhausted at the termination of the culture, the radioactivity was almost evenly distributed in u.v. absorbing materials in the P-I peak. The specific activity of the P-II peak was considerably lower than that of P-I RNA in the incorporation time here studied (Fig. 1). In Exp. B, however, the situation was reversed. Almost 85% of the radioactivity due to added [^{14}C]-adenine was in the P-II peak, only a small amount of radioactivity being detectable in the P-I peak (Fig. 2). This was also the case for the Rb of Exp. C (Fig. 3). This result indicates that azaguanine almost completely inhibits the formation of normal ribosomes and induces

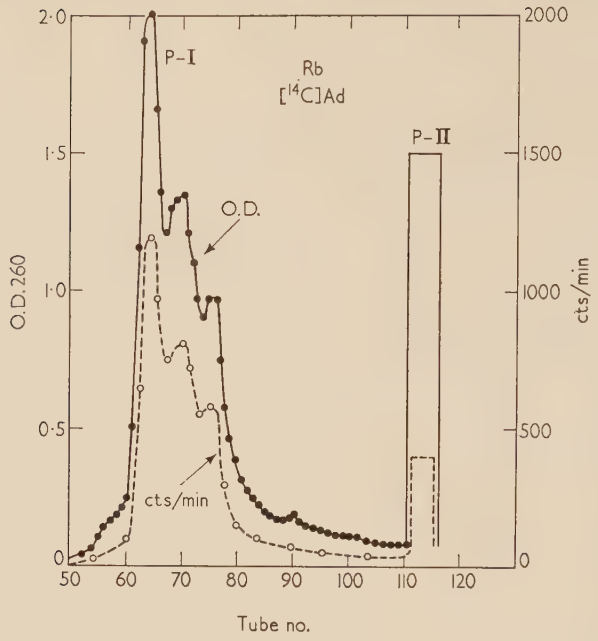


Fig. 1

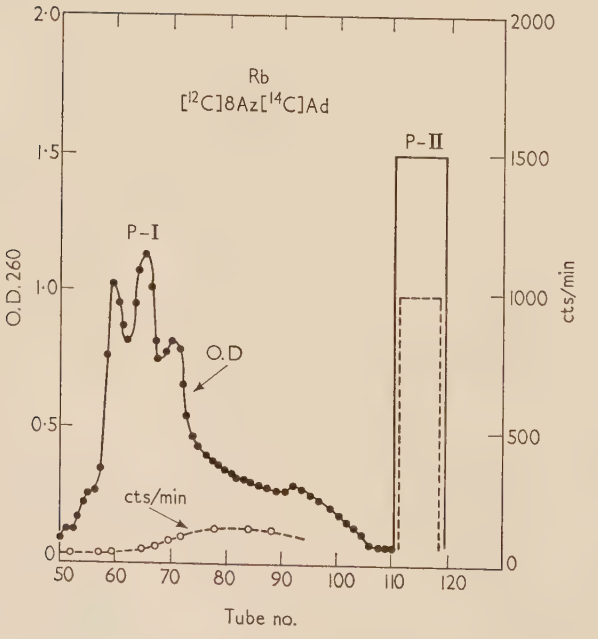


Fig. 2.

Figs. 1 to 6. Elution patterns on a DEAE column of the 105,000 *g* pellet and supernatant fraction of cell extract which was prepared after growing cells were exposed to [14C]-adenine (Figs. 1 and 4), [12C]-azaguanine plus [14C]-adenine (Figs. 2 and 5) or [14C]-azaguanine (Figs. 3 and 6) 6 ml./tube/ 15 min of effluent were collected at room temperature. Only the nucleic acid and nucleoprotein region of the pattern is shown. For further details see the text.

accumulation of the abnormal ribosomal RNA precursor *in the ribosome fraction.* That the RNA in the P-II peak may very well be the precursor of the ribosome RNA

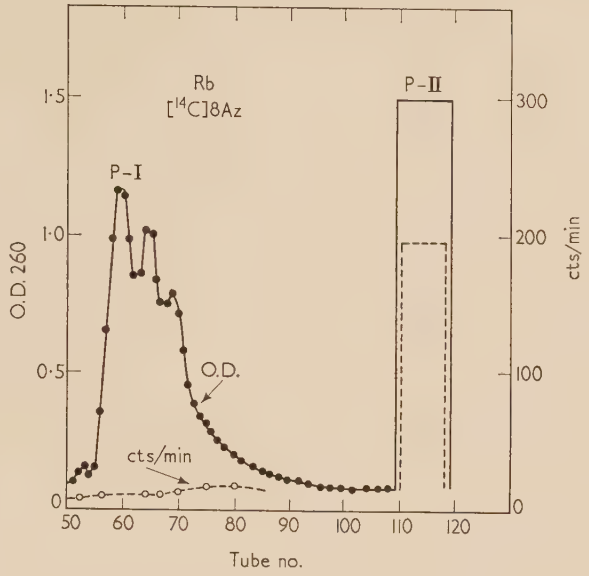


Fig. 3.

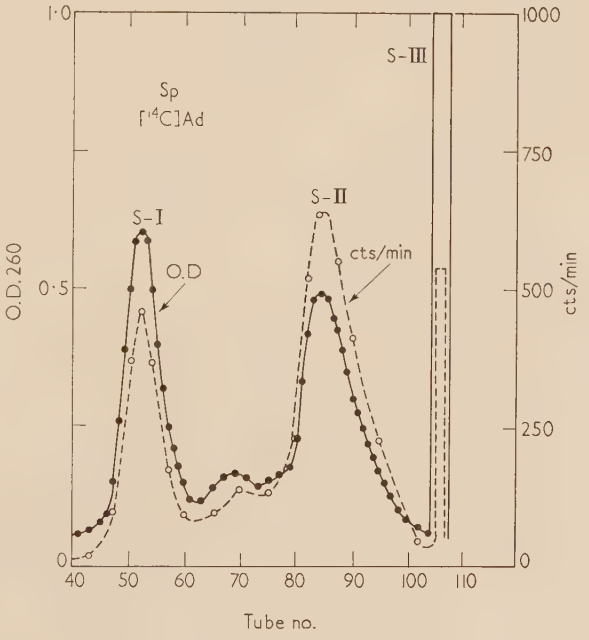


Fig. 4.

was suggested by (1) the above incorporation data, and (2) direct nucleotide composition analyses. The RNA in the P-II peak derived from the control cells has the same composition as the ribosome RNA. In the corresponding RNA from the culture fed

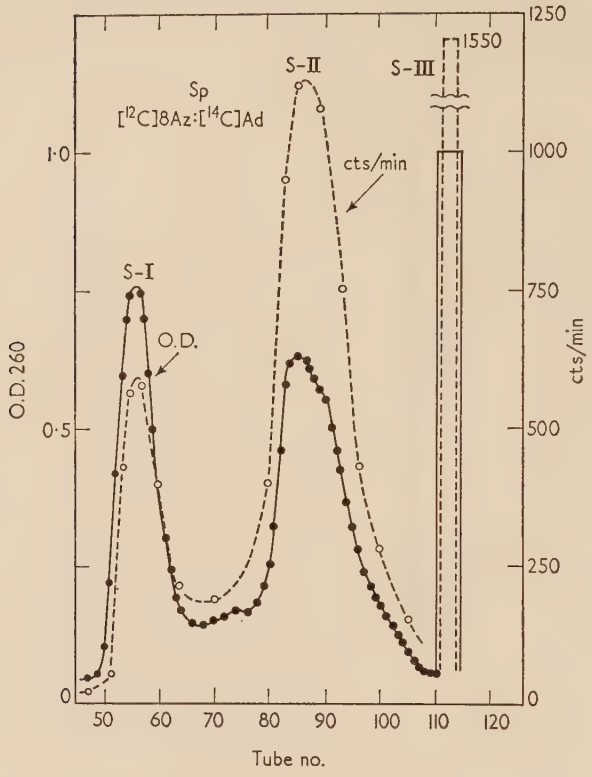

Sp
[12C]8Az:[14C]Ad
S-I
S-II
S-III
1550
cts/min
O.D.
O.D.260
cts/min
1·0
0·5
0
1250
1000
750
500
250
0
50 60 70 80 90 100 110 120
Tube no.

FIG. 5.

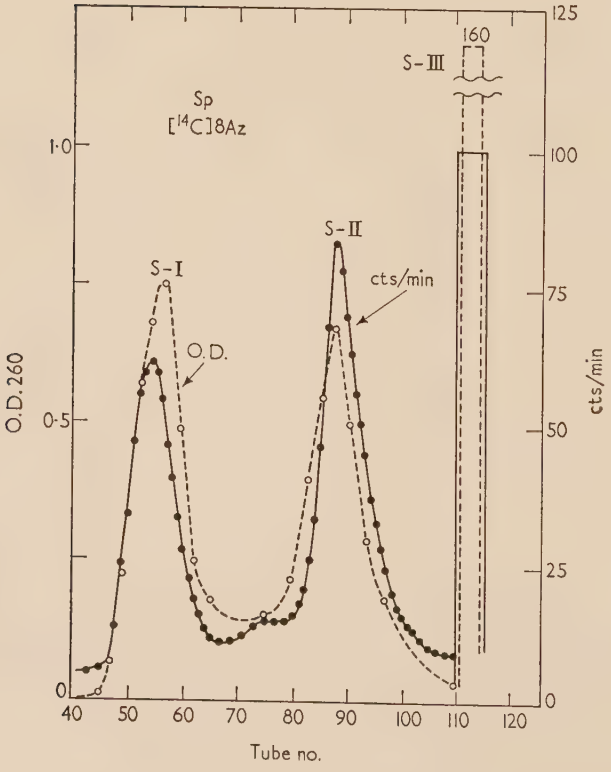

Sp
[14C]8Az
S-I
S-II
S-III
160
cts/min
O.D.
O.D.260
cts/min
1·0
0·5
0
125
100
75
50
25
0
40 50 60 70 80 90 100 110 120
Tube no.

FIG. 6.

with azaguanine, about 40% of the guanine was replaced by azaguanine, otherwise it had the same nucleotide composition as the normal ribosome RNA.

The fractionation pattern of the Sp on the DEAE column showed three nucleic acid peaks (Figs. 4, 5, 6). The S-I peak is S-RNA. The S-II peak is mainly DNA, because pretreatment of the Sp with deoxyribonuclease completely removed this peak without affecting the S-I peak. The S-III peak which was eluted with alkali was composed mainly of RNA but contained some DNA. In the azaguanine-containing culture an increase of the S-III peak relative to the S-I and S-II peaks was noted. The specific activity of S-RNA in Exp. 2 was not very much different from that of the control S-RNA, indicating that S-RNA synthesis was not very much affected by azaguanine, although a considerable amount of azaguanine was incorporated into S-RNA (*cf.* Fig. 6, S-I peak). The result is in accord with the data previously reported (Otaka, 1960). An interesting situation was seen in the S-II peak ("DNA" peak). The specific activity of S-II RNA in the control cells was a little higher than that of S-I RNA (Fig. 4). It was however seen in Exp. B that the activity of the peak was almost twice that of S-I or S-II in the control (Fig. 5). Furthermore in Exp. C considerable radioactivity was incorporated into "DNA" material (Fig. 6). Since azaguanine at this concentration is known to reduce the rate of DNA synthesis (Chantrenne & Devreux, 1960), it is not likely that the specific activity of DNA in Exp. B is higher than that of DNA in Exp. A. 42% of radioactivity in the S-II peak of Exp. B was made acid soluble by the ribonuclease treatment. These results strongly suggest that a large part of the activity in the S-II peak in Exp B is due to RNA. This suggestion may be further strengthened by the fact that while [^{14}C]-azaguanine was considerably incorporated into the S-II peak (Fig. 6) azaguanine is reported not to be appreciably taken up by DNA (Smith & Matthews, 1957). At any rate the inseparable behaviour of this "RNA" with the DNA on a DEAE column may indicate that the S-II peak consists of a complex of RNA and DNA. The presence of a small amount of radioactivity in the S-II peak due to RNA in the control experiment was also shown by the appearance of acid soluble radioactivity after treatment of the peak with ribonuclease.

In Exp. B, the specific activities of the RNA in different peaks (excepting S-RNA) were reduced in the direction of S-II > S-III > P-II > P-I; in the control culture the direction was P-I > P-II > S-III (> S-II?). This and the other data presented in this paper suggest that RNA is synthesized first on DNA and then transported as the precursor of ribosomal RNA onto old ribosomal particles, where it combines with protein to form new particles. Since azaguanine inhibits the synthesis of protein, presumably including ribosomal protein, the azaguanine-containing precursor RNA which has been synthesized has been merely deposited on the old ribosomal particles as suggested by the decreasing specific activity: P-II → P-I in Figs. 2 and 3. In this respect, it resembles the case of T2 bacteriophage infection of *Escherichia coli* (Nomura, Hall & Spiegelman, 1960), which is now being reinvestigated by us with the DEAE system described here.

Biological Institute
and Institute for Molecular Biology
Faculty of Science
Nagoya University
Chikusa-ku
Nagoya, Japan

E. Otaka
S. Osawa
Y. Oota

Received 27 March 1961

REFERENCES

Bolton, E. T., Britten, R. J., Cowie, D. B. & Roberts, R. B. (1958). *Publ. Carneg. Instn.* 57, 127.
Chantrenne, H. & Devreux, S. (1960). *Biochim. biophys. Acta*, **39**, 486.
Nomura, M., Hall, B. D. & Spiegelman, S. (1960). *J. Mol. Biol.* **2**, 306.
Otaka, E. (1960). *Exp. Cell. Res.* **21**, 229.
Smith, J. D. & Matthews, R. E. F. (1957). *Biochem. J.* **66**, 323.

J. Mol. Biol. (1961) **3**, 699–700

The Electron Microscopy of Phage DNA Molecules with Denatured Regions

The denaturation of DNA has been thought to be an "all-or-none" phenomenon. That is, either a single molecule retains its secondary structure completely, or it is rapidly converted into the "denatured" form which is pictured as randomly coiled polynucleotide chains (Doty, Marmur & Sueoka, 1959). When T2 or T4 bacteriophage DNA is thermally denatured, the extinction coefficient increases quite sharply within a two or three degree range about 83°C in normal saline (Doty *et al.*, 1959; Berns & Thomas, 1961). While this transition is sharp as the temperature increases, it is by no means a discontinuity.

With the finding that there is only a single DNA molecule in T2 or T4 phage (Thomas & Berns, 1961; Rubenstein, Thomas & Hershey, 1961; Beer, Naimark & Hershey, 1961), the "all-or-none" hypothesis leads to a difficulty: why should some molecules be completely denatured at lower temperatures than others, in spite of the genetic uniformity of the phage particles? This difficulty has led us to prefer the idea that a single molecule can be partially denatured; that is it can have regions of duplex and regions of single polynucleotide chains. We further supposed that this condition could exist in solution for a measurable time at the transition temperature.

It is now well established that, under favorable circumstances, denatured DNA can recover a significant portion of its secondary structure (Doty, Marmur, Eigner & Schildkraut, 1960; Marmur & Lane, 1960). Those regions of unorganized polynucleotide chains which are united to regions of completely ordered material might be expected to have a strong tendency to recover, because the register of the polynucleotide chains would be assured by the connecting intact regions which had never been separated. On cooling, these regions might be repaired and thus escape detection. This re-formation of hydrogen-bonded structure can be prevented to a large extent with formaldehyde. The formaldehyde reacts with the individual nucleotides and presumably interferes with the hydrogen bonding sites; as a result the polynucleotide chains cannot reassemble, as judged by optical density. In this way it has been possible to "trap" molecules in intermediate states of denaturation and to band them in CsCl (Berns & Thomas, 1961).

DNA solutions (15μg/ml. in 0·10 M-NaCl, 0·032 M-phosphate, 1% CH_2O, pH 7) were heated to various temperatures for 3 min and then chilled in ice water. The increase in optical density was measured and in some cases the solutions were banded and the mean density measured. After cooling they were rapidly dialysed to remove the salts against cold water containing 1% CH_2O. Small drops were spread in very thin films on cleaved mica with the side of a glass rod. After drying, pseudo-replicas were taken as described by Hall (1956).

Unheated DNA shows the now familiar picture of strands of DNA having a diameter of 20 $\pm$ 10 Å. Plate I(a) illustrates the kind of pictures that can be obtained by heating to 75°C. This temperature results in a 25% increase in extinction coefficient. Whereas total denaturation causes the apparent mean density of the DNA banding in CsCl to increase by 0·015 g/cm^3, this low temperature results in an increase

of approximately 0·0027 to 0·003 g/cm^3 or 18 to 20%. The molecule appears to have "puddles" of denatured polynucleotide chains united to intact segments. Heating to 100°C results in the disappearance of the native regions, and the only structures seen are the puddles of polynucleotide chains (Plate I(b)).

It is quite difficult to make an absolute correlation between extinction coefficient, density in CsCl and the number of denatured regions per unit length, mainly because of the lack of control of such important variables as salt concentration and formaldehyde content during the mounting of the molecules on mica. Some pictures showing loops have been obtained by heating to 83°C while fully denatured material has been seen at temperatures as low as 78·6°C. Therefore, it has not been possible to make a meaningful sequence of pictures taken on material treated at increasing temperatures.

It is of some importance to know whether the puddles seen in Plate I are located at random places or whether they mark specific "weak" points in the DNA molecule. To establish this will require visualizing longer segments of the molecule or the chemical identification of the denatured regions.

The competent assistance of Miss Doris Naimark is gratefully acknowledged, We wish to thank Mr. Kenneth Berns for his participation in this work. This work was supported by NSF grants G 7595 and G 10726, and NIH grant E-3233.

The Thomas C. Jenkins Department of Biophysics
The Johns Hopkins University
Baltimore, U.S.A.

Michael Beer
Charles A. Thomas, Jr.

Received 21 April, 1961

REFERENCES

Beer, M., Naimark, D. & Hershey, A. D. (1961). In preparation.
Berns, K. I. & Thomas, C. A., Jr. (1961). *J. Mol. Biol.* **3**, 289.
Doty, P., Marmur, J. & Sueoka, N. (1959). *Brookhaven Symposia in Biology* No. 12. p. 1.
Doty, P., Marmur, J., Eigner, J. & Schildkraut, C. (1960). *Proc. Nat. Acad. Sci., Wash.* **46**, 461.
Hall, C. E. (1956). *J. Biophys. Biochem. Cytol.* **2**, 625.
Marmur, J. & Lane, D. (1960). *Proc. Nat. Acad. Sci. Wash.* **46**, 461.
Rubenstein, I., Thomas, C. A. & Hershey, A. D. (1961). *Proc. Nat. Acad. Sci., Wash.* in the press.
Thomas, C. A. & Berns, K. I. (1961). *J. Mol. Biol.* **3**, 277.

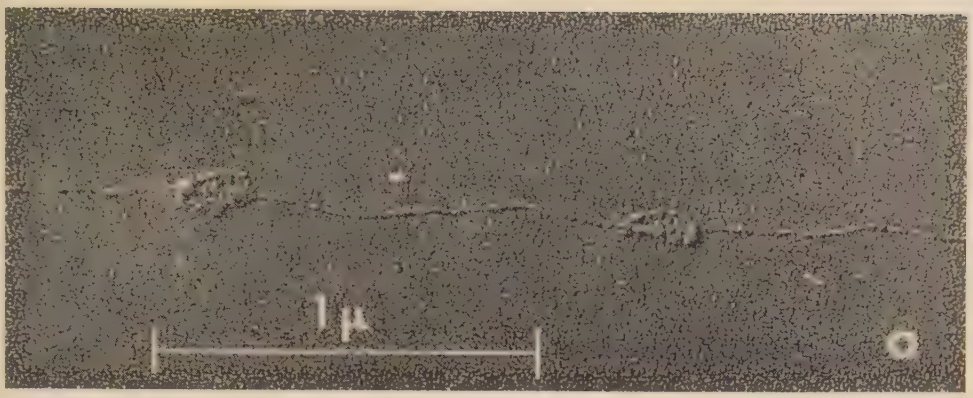


PLATE I(a). DNA from preparation which has been heated to 75°C in the presence of formaldehyde. Puddles along fibril are believed to be denatured DNA. Magnification × 50,000.

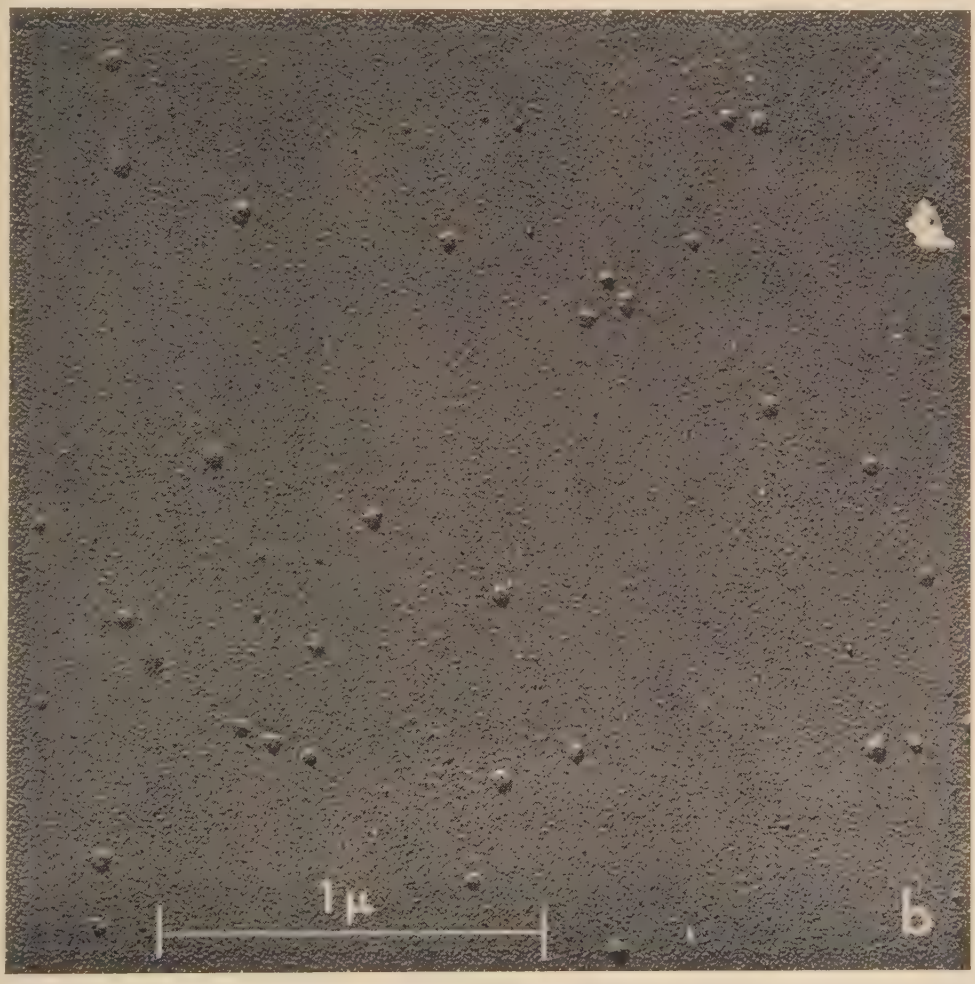


PLATE I(b). DNA from preparation which has been heated to 100°C in the presence of formaldehyde. Magnification × 50,000.

[*To face page* 700

J. Mol. Biol. (1961) **3**, 701–702

Surface Structure of Tobacco Mosaic Virus†

Purified preparations of tobacco mosaic virus (TMV) examined by electron microscopy appear to consist of rods about 150 Å in diameter having various lengths. When sufficient care is exercised in the preparation, rods with a length of 3000 Å predominate (Williams & Steere, 1951), and these are considered to be the infectious units of mosaic disease (Sigureirsson & Stanley, 1947). Each of these rods seems to consist of a single polyribonucleotide chain surrounded by a "coat" of protein (Hart, 1958). This coat, according to recent results and interpretations of X-ray diffraction studies (Klug & Caspar, 1960), is made of some 2130 identical globular subunits packed in a helical array of pitch 23 Å. The rod is thought to be encircled by a pronounced helical groove, which contains the RNA chain and demarcates adjacent turns of the protein helix (Franklin & Klug, 1956), though such a feature has not appeared, until now, in electron micrographs of TMV.

The spacings of the protein subunits, as determined by X-ray diffraction, are large enough to be resolvable by commercially available electron microscopes; however the usual specimen preparation techniques do not render such fine detail visible. There have been isolated reports of regular cross-striations, of various spacings, seen on TMV particles treated by the metal shadow-casting technique, but the spacings did not even approximate those predicted by X-ray diffraction, and it is believed that the striated appearance in such cases was due to astigmatism or specimen movement during the exposure of the micrograph. The highest quality electron micrographs of shadowed TMV published to date do not reveal any regular surface structure (Williams, 1952). Tobacco mosaic virus protein rods (without nucleic acid) when "negatively stained" with potassium phosphotungstate clearly showed cross-striations spaced about 23 Å apart (Nixon & Woods, 1960); yet when the same treatment was applied to the complete virus particles, no regular surface features were visible. This latter result is rather surprising, since nearly identical arrangements of the protein subunits are indicated by the X-ray diffraction patterns of the virus and the protein rods (Franklin, 1955). One possible explanation is that in the protein rods the phosphotungstate occupies sites "vacated" by the nucleic acid, thus staining the helical groove and rendering it visible.

In the present study, purified TMV suspended in Pyrex-distilled water was sprayed onto nickel mesh grids coated with collodion and carbon. The grids were then placed under the bell jar of a vacuum evaporation apparatus (Kinney Mfg. Co., model SC-3, modified to include a thimble-type cold trap between the oil diffusion pump and the main line to the bell jar) and shadowed with tungsten, which was evaporated by passing about 30 amperes through a straight piece of the metal wire 0·48 mm in diameter. Shadowing incidence angles of about 45° were used.

Specimens were examined and their micrographs taken with a Siemens Elmiskop I electron microscope. The instrument was used without an objective aperture, with the two-lens condensing system, and with the control settings appropriate for a nominal

† This research was supported by a grant from the National Institute of Arthritis and Metabolic Diseases, U.S. Public Health Service.

magnification on the photographic plate of 160,000, a value high enough to eliminate possible confusion on the micrographs between photographic grain and specimen detail.

Plate I is a selected field showing a number of rods in a TMV preparation shadowed with tungsten. The deposition was light, and in preparing this figure the contrast had to be greatly enhanced photographically over that in the original plate. Indications of a regular arrangement of subunits may be seen on some of the particles. A few rods have regularly spaced striations almost normal to their axes over portions of their lengths. Such striations appear on rods lying in various directions and are therefore not attributable to astigmatism or specimen movement, which can produce streaks of only one direction in any one micrograph. Where the striations are regular, they are spaced very nearly 23 Å apart, and it is believed that they represent individual turns of the protein helix. The striations are not smooth, but seem to consist of rows of "bumps," which may be the globular protein subunits of about 18,000 molecular weight.

It is not understood why tungsten shadowing permits observation of TMV helical structure on some, but not all, of the rod surfaces. Perhaps there is enough local variation in shadow film quality to account for this inconsistency. Alternatively, it may be proposed that the shadow film is not at fault but that the protein subunits tend to become disordered from their regular array while drying, and that the various degrees of disorder seen in any field are the results of local surface tension forces.

Department of Zoology and the Adolphus Busch III Laboratory of Molecular Biology
Washington University
St. Louis, Missouri
U.S.A.

ROGER G. HART

Received 21 June 1961

REFERENCES

Franklin, R. E. (1955). *Biochim. biophys. Acta*, **18**, 313.
Franklin, R. E. & Klug, A. (1956). *Biochim. biophys. Acta*, **19**, 403.
Hart, R. G. (1958). *Biochim. biophys. Acta*, **28**, 457.
Klug, A. & Caspar, D. L. D. (1960). *Advanc. Virus Res.* **7**, 225.
Nixon, H. L. & Woods, R. D. (1960). *Virology*, **10**, 158.
Sigureirsson, T. & Stanley, W. M. (1947). *Phytopathology*, **37**, 20.
Williams, R. C. (1952). *Biochim. biophys. Acta*, **8**, 227.
Williams, R. C. & Steere, R. L. (1951). *J. Amer. Chem. Soc.* **73**, 2057.

PLATE I. A tobacco mosaic virus preparation shadowed with tungsten at about 45° incidence. Some of the rods exhibit surface detail indicative of their helical structure. × 350,000.

[*To face page* 702

J. Mol. Biol. (1961) **3**, 703–704

A Temperature-Sensitive Regulatory System†

We would like to report the finding of a strain of *Escherichia coli* that has a temperature-sensitive mechanism for the regulation of the rate of synthesis of the enzyme β-galactosidase.

Normally, the synthesis of this enzyme is blocked by the presence of a specific repressor substance and synthesis occurs only in the presence of a specific inducer (Pardee, Jacob & Monod, 1959). There are mutants, called constitutive, which are able to make β-galactosidase even in the absence of inducer apparently because they are unable to make the repressor substance.

A variant of the methionine-requiring Hfr strain E102 (Horiuchi, Tomizawa & Novick, 1961) has been obtained which, to a first approximation, behaves like an inducible strain at lower temperatures of incubation and like a constitutive mutant

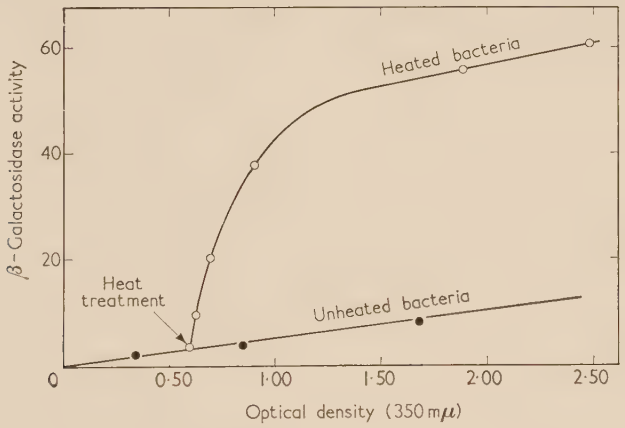


Fig. 1. Demonstration of the thermolability of the repression system in E103. The procedure is described in the text. Enzyme is measured in units of 0·89 mμ mole of ONPG hydrolysed per minute at 28°C at pH 7 (Horiuchi *et al.*, 1961). An optical density of 0·100 corresponds to 3×10^7 bacteria/ml., or about 1 μg/ml. bacterial protein nitrogen.

at higher temperatures. This strain (E103) in the absence of inducer makes β-galactosidase at 14°C at about 0·5% of the maximum rate while at 43·5°C it makes the enzyme at about 40% of the maximum rate. In comparison, the wild type strain in the absence of inducer makes β-galactosidase at less than 0·1% of the maximum rate at both temperatures. When inducer is present the rate of synthesis is at maximum at all temperatures both for wild type bacteria (E102) and for E103. The enzyme produced in the absence of inducer cannot be attributed to the presence of an admixture of constitutive bacteria.

An experiment illustrating the thermal lability of the regulatory system in E103 is shown in the figure. Here E103 bacteria were grown at 14°C in lactate medium with

† This work was supported in part by United States Public Health Service research grant no. E2807, and training grant no. 2A-14 PB and in part by Office of Naval Research contract no. 2771.

methionine (doubling time of 12 hr). At one point they were washed and resuspended in phosphate buffer medium in which they could not grow because of a lack of a carbon source and methionine. The bacteria were heated in this buffer medium at 45°C for 30 min with little loss in viability and then returned to cold lactate medium with methionine for further cultivation at 14°C. It can be seen from the figure that following heating there was an abrupt rise in the rate of β-galactosidase synthesis. The rate remained high for about a generation and then fell to the low rate characteristic of this strain at 14°C.

Experiments of this kind were performed with varied times of heating and it was observed that 15 min at 45°C is sufficient to produce the maximum effect.

The behaviour of E103 following heating could be the result of the presence in this mutant of a thermolabile repressor substance. On this basis, after return to 14°C the period of time of about a generation required before the rate of synthesis falls to its normal value would be the time required for sufficient newly-made repressor to accumulate, much as observed in experiments with appropriate newly-formed zygotes (Pardee *et al.*, 1959).

However, there are several alternative hypotheses which could account for the behaviour of E103, and two of these will be discussed. It might be that in E103 an enzyme needed for repressor formation is thermolabile. In this case, to explain the present experiments, it is necessary to assume that the amount of repressor present at any time is proportional to the level of this enzyme, which would imply that the repressor is turned over very rapidly.

Another possible explanation is that the effect of higher temperature on E103 is to produce a modification of the operator gene, the postulated site of action of repressor (Jacob, Perrin, Sanchez & Monod, 1960; Jacob & Monod, 1961). Such an explanation would require not only the assumption that heating alters the operator so that it is unaffected by repressor but also the assumption that this alteration is slowly reversed at 14°C.

Institute of Molecular Biology
Eugene
Oregon, U.S.A.

Tadao Horiuchi†
Sumiko Horiuchi†
Aaron Novick

Received 6 June 1961

REFERENCES

Horiuchi, T., Tomizawa, J. & Novick, A. (1961). In preparation.
Jacob, F. & Monod, J. (1961). *J. Mol. Biol.* **3**, 318.
Jacob, F., Perrin, D., Sanchez, C. & Monod, J. (1960). *C. R. Acad. Sci., Paris*, **250**, 1727.
Pardee, A., Jacob, F. & Monod, J. (1959). *J. Mol. Biol.* **1**, 165.

† On leave from the National Institute of Health of Japan, Tokyo, Japan.

J. Mol. Biol. (1961) **3**, 705–706

Sidechain-Sidechain Interaction between the α-Helices of L- and D-Polypeptide Chains

In the course of our work on the structures of synthetic polypeptides, we have observed the infrared absorption spectrum of a film cast from a dioxane solution of a 50 : 50 mixture of high molecular-weight poly-γ-benzyl-L-glutamate and poly-γ-benzyl-D-glutamate. The observed spectrum, given in Fig. 1, is somewhat different from the spectrum given in a previous paper (Tsuboi, 1961) of a film of the pure L- or pure D-polypeptide. A slightly more detailed examination of the two spectra

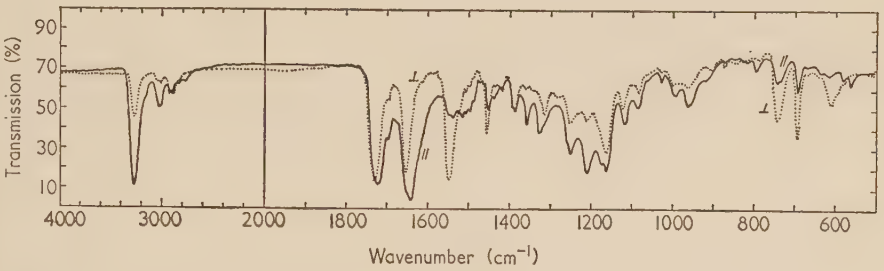


FIG. 1. Infrared absorption spectrum of an oriented film cast from a dioxane solution of a 50 : 50 mixture of poly-γ-benzyl-L-glutamate and poly-γ-benzyl-D-glutamate. Full line, electric vector of the incident radiation parallel to the fiber axis; dotted line, electric vector of the incident radiation perpendicular to the fiber axis.

shows that the differences are mostly in the relative intensities and dichroic properties of the side-chain bands. The positions, intensities, and dichroic properties of the NH stretching, amide I, amide II, amide III and amide V bands at about 3290, 1650, 1550, 1315, and 613† cm^{-1} respectively, were found to be similar.‡ While, for example, the dichroic ratios ($\perp/\parallel$) of the phenyl B_2 bands at about 750 and 700 cm^{-1} greatly increase, the band shape of the phenyl B_2 band at about 750 cm^{-1} and of the ester C—O band at about 1160 cm^{-1} are altered, and the relative intensities of the two bands at 1360 and 1210 cm^{-1} (which are also probably to be ascribed to side-chain vibrations) greatly increase on going from the pure L (or D) film to the D-L mixture. From these facts, it may be concluded that there is an appreciable interaction between the side-chains of adjacent polypeptide chains and that the interaction between the D- and L-chains is somewhat different from that between two L-chains or between two D-chains.

An X-ray diffraction study has also been made of the mixture of poly-γ-benzyl-L-glutamate and poly-γ-benzyl-D-glutamate. From a concentrated, viscous dioxane solution of a 50 : 50 mixture of the D- and L-polypeptides, a fiber was drawn out. Its fiber diagram was found to be somewhat different from that of the fiber of pure L- or pure D-polypeptide. The most marked difference is the following: on the latter a

† The perpendicular band at 613 cm^{-1} was assigned to the amide VI band in a previous paper (Tsuboi, 1961). However, since Miyazawa (1961) found that this band is shifted to 458 cm^{-1} by N-deuteration, it is better to assign it to the amide V band.

‡ In addition to these amide bands, those at 799, 643, and 563 cm^{-1} are assignable to main chain vibrations characteristic of the α-helix (Miyazawa, 1961; Tsuboi, 1961). All these bands are observed for the D-L mixture film at the same positions with the same intensities and dichroic properties as those observed for the pure L-film.

strong meridional arc of 5·3 Å is observed while in the former this is very faint; instead, in the former a strong meridional arc of 10·4 Å is observed which is practically absent in the latter. In both the fiber diagrams a meridional reflection of 1·5 Å is observed, and therefore the chain conformation is considered to be that of the α-helix in both fibers. The 5·3 Å meridional reflection corresponds to nearly one pitch of the α-helix and the 10·4 Å reflection to nearly two pitches. One of the possible explanations of the appearance of this 10·4 Å fiber period in the D-L mixture fiber is given by assuming a special sidechain-sidechain interaction between L- and D-chains running antiparallel to each other and forming a pair with a center of symmetry (see Fig. 2).

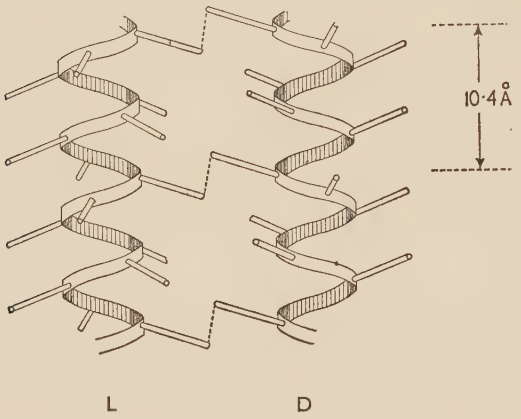


Fig. 2. Schematic view of a possible pairing of α-helices of poly-γ-benzyl-L-glutamate and poly-γ-benzyl-D-glutamate.

If, in a pitch of α-helix, 3·5 residues (instead of 3·6 residues) are involved, such a sidechain-sidechain interaction may take place once every two pitches of the helix. The pair of α-helices with such a sidechain-sidechain interaction could be constructed by only a slight deformation of a "natural" 5-turn 18-residue α-helix which has 3·6 residues per turn. This pair of α-helices (in a crystalline arrangement) would show the 10·4 Å fiber period, just as is actually observed.

A more detailed study on these lines is now in progress, and the results will be published later.

The authors wish to express their thanks to Professor T. Shimanouchi and Dr. H. Oeda for their interest and encouragement.

Faculty of Pharmaceutical Sciences
University of Tokyo
Bunkyo-ku, Tokyo, Japan
Masamichi Tsuboi

Faculty of Science
Ochanomizu University
Bunkyo-ku, Tokyo, Japan
Akiyoshi Wada

Central Research Laboratory
Ajinomoto Company
Kawasaki, Kanagawa, Japan
Nobuya Nagashima

Received 6 June 1961

REFERENCES

Miyazawa, T. (1961). 1st International Symposium on Poly-α-Amino Acid, University of Wisconsin, U.S.A.

Tsuboi, M. (1961). *J. Polymer Sci.* In the press.

J. Mol. Biol. (1961) **3**, 707–708

Haemoglobin I_2, a Further Variant of Haemoglobin A_2

Recent investigations have shown that the α-chains of haemoglobins A and A_2 are similar (Müller & Jonxis, 1960; Huehns & Shooter, 1961*a*,*b*). The latter authors have suggested that the final stage in the synthesis of haemoglobin $A_2(\alpha_2\,\delta_2^{A_2})$ is the combination of independently synthesized α_2^A and $\delta_2^{A_2}$ subunits, the α_2^A subunits for the formation of both haemoglobins being derived from a common pool. They pointed out that on this hypothesis an individual who is synthesizing abnormal α_2^X subunits will produce not only a variant of haemoglobin A, $\alpha_2^X\beta_2^A$, but also the corresponding variant of haemoglobin A_2, $\alpha_2^X\delta_2^{A_2}$. In support of this they cited the occurrence, in a number of individuals carrying haemoglobin G ($\alpha_2^G\beta_2^A$), of a second minor haemoglobin, haemoglobin G_2, which had a similar electrophoretic mobility to the artificially produced species, $\alpha_2^G\delta_2^{A_2}$. Further evidence in favour of this hypothesis comes from the finding, reported here, of both haemoglobin A_2 and a second minor haemoglobin, haemoglobin I_2 of probable composition $\alpha_2^I\delta_2^{A_2}$, in an individual carrying the haemoglobins A and I.

In the haemolysate of the individual previously reported (Atwater, Schwartz, Erslev, Montgomery & Tocantins, 1960), haemoglobin I formed 80% of the total haemoglobin. On starch block electrophoresis the haemoglobin I separated readily from the second, slower zone which migrated in the position of haemoglobin A. Only small amounts of haemoglobin A_2 were detected. Electrophoretic analysis in starch gel, as previously described (Gammack, Huehns, Shooter & Gerald, 1960), confirmed that the relative amount of haemoglobin A_2 was much less than in normal adult haemolysate. Two other zones corresponding to haemoglobins I and A were also observed (Plate I).

Haemoglobin I is known to have abnormal α_2 subunits (Ingram, 1959); the α-chain of haemoglobin I differs from that of haemoglobin A by the replacement of a lysyl residue by an aspartyl residue and has the formula $\alpha_2^{16\text{Asp}}\beta_2^A$ (Murayama, 1959). At the pH of the starch gel analyses, haemoglobin I therefore differs in net charge from haemoglobin A by − 4. Moreover, haemoglobin A_2 differs in net charge from haemoglobin A by about + 4, these extra positive charges residing in the $\delta_2^{A_2}$ subunit. Thus, the haemoglobin species $\alpha^I\delta_2^{A_2}$ would have the same net charge as haemoglobin A. This conclusion has been verified by means of a recombination experiment (Itano & Robinson, 1959; Gammack *et al.*, 1960) using isolated haemoglobin I and a purified preparation of haemoglobin A_2 from a normal adult (Huehns & Shooter, 1961*b*). When a mixture of haemoglobin I and haemoglobin A_2 was exposed to a pH of 4·7 for 96 hours and then dialysed to a neutral pH, conditions under which haemoglobins A_2 and G_{Ib}, $\alpha_2^G\beta_2^A$ (Gammack *et al.*, 1960), form significant amounts of the haemoglobin species $\alpha_2^G\delta_2^{A_2}$ (Huehns & Shooter, 1961*a*,*b*), analysis revealed the presence of a single new haemoglobin zone migrating in the position of haemoglobin A (Plate II). This result is consistent with the exchange of subunits between haemoglobins I and A_2, according to the following equation:

$$\alpha_2^I\beta_2^A + \alpha_2^A\delta_2^{A_2} \rightleftharpoons \alpha_2^A\beta_2^A + \alpha_2^I\delta_2^{A_2}$$

	Hb-I	Hb-A_2	Hb-A	Hb-I/A_2
Net charge difference from Hb-A	− 4	+ 4	0	0

The new haemoglobin zone would thus contain both haemoglobin A and the new species haemoglobin I/A_2, usually called haemoglobin I_2. It follows that naturally occurring haemoglobin I_2 would also migrate with haemoglobin A. Therefore, recombination of the isolated "haemoglobin A" fraction by itself would result in the formation of both haemoglobins I and A_2, as indicated by the reverse of the above reaction provided it contained haemoglobin A as well as haemoglobin I_2. The presence of haemoglobin A was confirmed by the production, in a recombination experiment with canine haemoglobin, of the same two new species that form in the recombination of haemoglobin A with canine haemoglobin (Plate III).

The analyses of the untreated and recombined "haemoglobin A" fractions show the formation of both haemoglobin A_2 (Plate IV(*a*)) and haemoglobin I (Plate IV(*b*)). It was not possible to demonstrate satisfactorily both these components in one electrophoretic analysis because the discontinuous buffer system gives relatively poor resolution of minor components which migrate faster than haemoglobin A, whilst the phosphate buffer system is insensitive for slow minor components. In parallel control experiments using the "haemoglobin A" fraction from the haemolysate of an individual carrying a haemoglobin variant, namely haemoglobin J_{Tr}, whose anodic mobility is also greater than that of haemoglobin A but which has abnormal β_2 subunits (Gammack, Huehns, Lehmann & Shooter, 1961), no new species were formed. The "haemoglobin A" fraction from the individual carrying Hb-I therefore contains haemoglobin A and a second haemoglobin with the probable molecular composition $\alpha_2^I\delta_2^{A_2}$, namely Hb-I_2. Moreover, the appearance of Hb-I_2 in this individual would account for the decreased amount of Hb-A_2 present.

The demonstration that an individual who carries haemoglobins A and I also has both haemoglobins A_2 and I_2 lends further support to the hypothesis that the final stage in the synthesis of haemoglobin A_2 is the combination of independently synthesized α_2^A and $\delta_2^{A_2}$ subunits and that the α_2^A subunit for haemoglobin A and A_2 formation, or the α_2^X subunits for the synthesis of α-chain variants of haemoglobins A and A_2, are derived from a common pool.

Cardeza Foundation
Jefferson Medical College and Hospital
Philadelphia, Penn., U.S.A.

JEAN ATWATER

Department of Biochemistry
University College
London W.C.1, England

E. R. HUEHNS†
E. M. SHOOTER

Received 15 June 1961

REFERENCES

Atwater, J., Schwartz, I. R., Erslev, A. J., Montgomery, T. L. & Tocantins, L. M. (1960). *New Engl. J. Med.* **263**, 1215.
Gammack, D. B., Huehns, E. R., Shooter, E. M. & Gerald, P. S. (1960). *J. Mol. Biol.* **2**, 372.
Gammack, D. B., Huehns, E. R., Lehmann, H. & Shooter, E. M. (1961). *Acta Genet.* **11**, 1.
Huehns, E. R. & Shooter, E. M. (1961*a*). *Nature*, **189**, 918.
Huehns, E. R. & Shooter, E. M. (1961*b*). *J. Mol. Biol.* **3**, 257.
Ingram, V. M. (1959). *Nature*, **183**, 1795.
Itano, H. A. & Robinson, E. (1959). *Nature*, **183**, 1799.
Müller, C. J. & Jonxis, J. H. P. (1960). *Nature*, **188**, 949.
Murayama, M. (1959). *Fed. Proc.* **19**, 78.

† Beit Memorial Fellow.

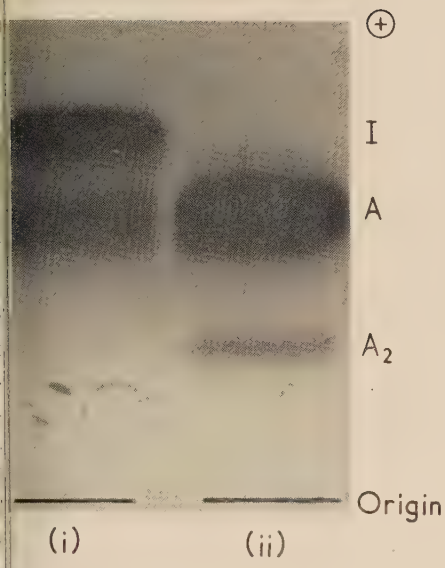


PLATE I. Starch gel analysis in the discontinuous system. Migration time 2 hr, o-dianisidine stain. (i) Haemolysate from individual carrying haemoglobin A and I, and (ii) normal adult haemolysate.

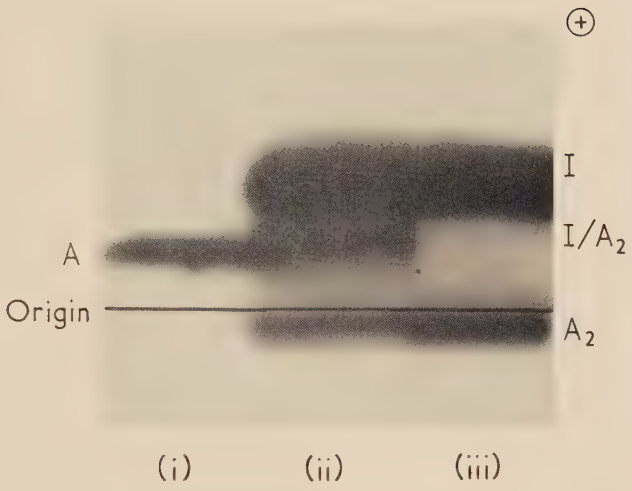


PLATE II. Starch gel analysis in the phosphate buffer system, pH 7·4. Migration time 2·5 hr, o-dianisidine stain. (i) Haemoglobin A marker. (ii) Mixture of haemoglobins A_2 and I dissociated and recombined. (iii) Untreated mixture of haemoglobins A_2 and I.

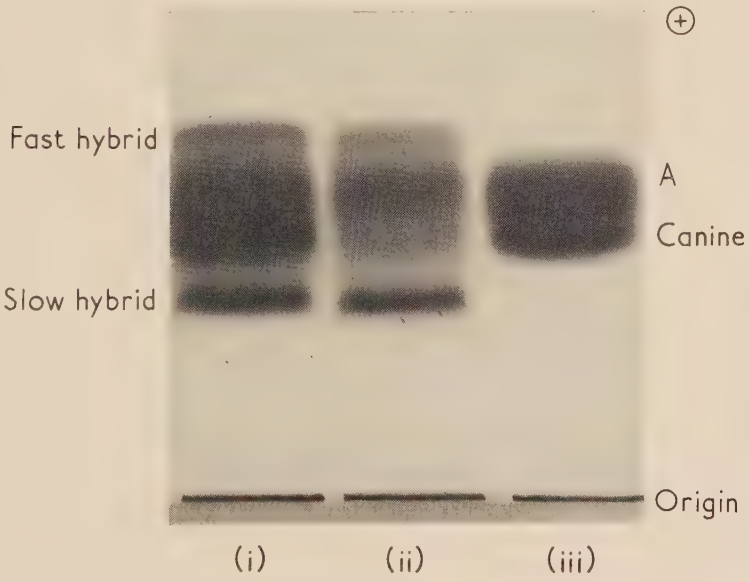
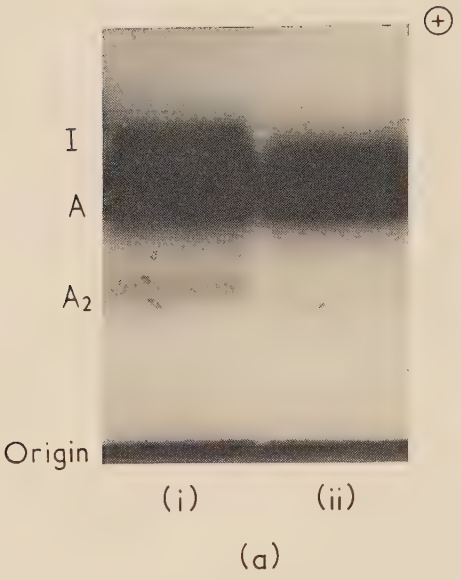


PLATE III. Starch gel analysis in the discontinuous buffer system. Migration time 2 hr, o-dianisidine stain. (i) Mixture of haemoglobin A (normal adult) and canine haemoglobin dissociated and recombined. (ii) Mixture of "haemoglobin A" fraction and canine haemoglobin dissociated and recombined. (iii) Untreated mixture of "haemoglobin A" fraction and canine haemoglobin.

[*To face page* 708



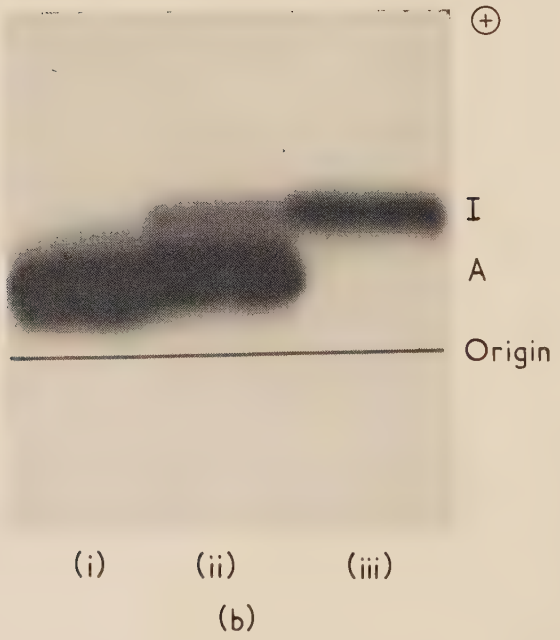


PLATE IV. (a) Starch gel analysis in the discontinuous system. Migration time 2 hr, *o*-dianisidine stain. (i) "Haemoglobin A" fraction dissociated and recombined, and (ii) untreated "haemoglobin A" fraction.

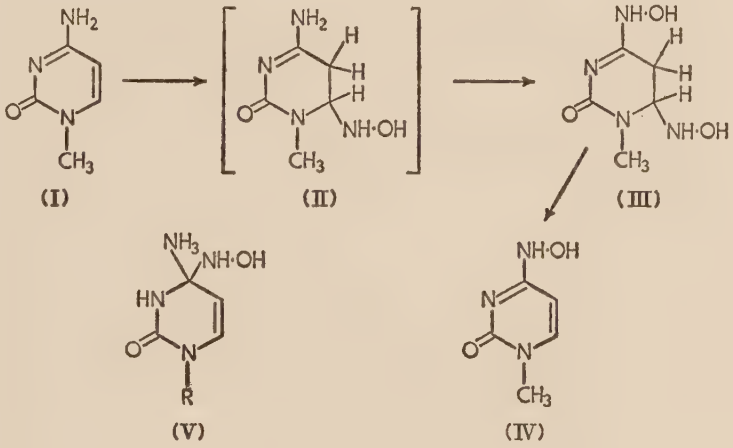
(b) Starch gel analysis in the phosphate buffer system. (i) Untreated "haemoglobin A" fraction. (ii) Dissociated and recombined "haemoglobin A" fraction. (iii) Haemoglobin I marker.

J. Mol. Biol. (1961) **3**, 709–710

The Reaction of Hydroxylamine with Cytosine and Related Compounds

The action of simple reagents on nucleic acids is of considerable interest from several points of view. Controlled degradation can give valuable information on nucleotide sequence. Related to this type of study is the field of chemical mutagenesis, and both would be materially advanced if reagents of higher specificity could be devised. Such specificity might relate either to the constituent purine or pyrimidine bases of the nucleic acid or to internucleotidic linkages between defined nucleoside residues.

Following an interest in hydrazine, which reacts only with the pyrimidine bases (Baron & Brown, 1955), we were led to study the action of hydroxylamine. In view of recent publications by Freese and his coworkers (Freese, Bautz-Freese & Bautz, 1961; Freese, Bautz & Bautz-Freese, 1961) on the mutagenic action of this reagent and the partial elucidation of the chemical aspects of this process, we record here some observations which help to clarify the latter.†



Herring sperm DNA was treated with hydroxylamine and samples removed at intervals. Hydrolysis of these by formic acid followed by chromatography revealed the progressive disappearance of cytosine and the appearance of a new base. A very small loss of thymine also occurred. It has been shown (Freese, Bautz & Bautz-Freese, 1961) that thymidine is apparently unaffected by hydroxylamine under conditions where cytosine and related compounds undergo reaction. We would note here, however, that reactions of the base residues in DNA are much slower than in the corresponding nucleosides and that a levelling effect on rates is observable.

As a convenient model substance for a study of the hydroxylamine reaction with cytosine and related residues in nucleic acids, 3-methylcytosine (I) was used. With anhydrous hydroxylamine or 5 N-hydroxylamine acetate, pH 6·6, it gave quantitatively a non-basic product $C_5H_{10}O_3N_4$ (mol. wt. 174; found 180). This had λ_{max} 220 mμ

†*Note added in proof:* Related observations have also been made by Zillig, Verwoerd & Kolhage (Abstracts of the symposium on Ribonucleic Acids and Polyphosphates, Strasbourg, July, 1961) and by H. Schuster (personal communication; *J. Mol. Biol.*, in the press).

indicating that the 4,5-double bond had become saturated, confirmed by a nuclear magnetic resonance (NMR) study in D_2O which showed the disappearance of the low field absorption due to the protons on C_4 and C_5 in (I) (Jardetsky & Jardetsky, 1960). The spectrum was consistent with the structure (III) showing, superimposed on the sharp resonance of the methyl group at $\tau = 6{\cdot}91$, a multiplet which could be interpreted as the resonances of the AB protons of an ABX system with chemical shifts about 6·85 and 7·25 p.p.m. A quartet of lines at $\tau = 5{\cdot}43$ corresponded, clearly, to the resonance of the X proton.

The product, with acid, gave a base, $C_5H_7O_2N_3$, isolated as its hydrochloride (equiv. wt. 177, found, 182) which is formulated as (IV) from the identity of its characteristic ultraviolet spectra (pH 6·5, λ_{max} 277, 235 mμ; pH O, λ_{max} 280 mμ) with those of N^6-hydroxycytidine (Fox *et al.*, 1959). (IV) formed a blue chelate with Fe^{3+} in solution, while (III) underwent oxidation by this ion. Our overall observations agree very closely with those of Freese & co-workers (1961) on cytosine and the derived nucleosides, but we differ in our interpretation of the structure of the hydroxylamine reaction product. Their structure (V) is unlikely on general grounds and is inconsistent with the NMR evidence. It may be noted, too, that the conversion of (III) to (IV), which proceeds slowly even at pH 6·6 and is rapid at low pH values, has its counterpart in the easy acid-catalysed dehydration of 4-hydroxy-4, 5-dihydro-uridine and -cytidine (Shugar, 1960). The exchange reaction, (II) → (III), analogous to the rapid base-catalysed hydrolysis of the amino-group of 4,5-dihydrodeoxycytidine (Green & Cohen, 1958) occurs at a rate similar to or faster than the addition, (I) → (II), since attempts to detect the presumed intermediate have so far failed. This point requires further study as it may be important to an understanding of the mutagenic action of hydroxylamine. Related too, is the question of the tautomeric equilibria in the compounds formulated here as (III) and (IV) which we are studying.

Hydroxylamine reacts in a quite analogous manner with cytosine (at about 5 times the rate) and with cytidine and deoxycytidine to give isolable products corresponding to (III).

Satisfactory analyses on all relevant compounds have been obtained and full details of the work will be published elsewhere.

University Chemical Laboratory
Lensfield Road
Cambridge

D. M. Brown
P. Schell

Received 5 July 1961

REFERENCES

Baron, F. & Brown, D. M. (1955). *J. Chem. Soc.* 2855.
Fox, J. J., Van Praag, D., Wempen, I., Doerr, I. L., Cheong, L., Knoll, J. E., Eidinoff, M. L., Bendich, A. & Brown, G. B. (1959). *J. Amer. Chem. Soc.* **81**, 178.
Freese, E., Bautz, E. & Bautz-Freese, E. (1961). *Proc. Nat. Acad. Sci., Wash.* **47**, 845.
Freese, E., Bautz-Freese, E., & Bautz, E. (1961). *J. Mol. Biol.* **3**, 133.
Green, M. & Cohen, S. S. (1958). *J. Biol. Chem.* **228**, 601.
Jardetsky, C. D. & Jardetsky, O. (1960). *J. Amer. Chem. Soc.* **82**, 222.
Shugar, D. (1960). In *The Nucleic Acids*, ed. by E. Chargaff & J. N. Davidson, Vol. 3. p. 73. New York: Academic Press.

INSTRUCTIONS TO AUTHORS

1. The *Journal of Molecular Biology* will publish papers on the nature, production and replication of biological structure at the molecular level, and its relation to function. Suitable subjects are sub-cellular organization: molecular genetics: structure and replication of viruses: molecular structure of muscle, nerve and other tissues: structure of proteins, nucleic acids, carbohydrates, lipids, etc., and their synthetic analogues, as investigated by X-rays, light absorption and other methods: problems of inter- and intra-molecular energy transfer. The Journal will not as a rule publish papers when the subject matter is outside the fields defined above; however, when the subject matter borders on molecular biology and is of special interest in relation to molecular biology such papers may be published when space permits.

2. The Journal will also publish Letters to the Editor (averaging 1000 words), and occasional review articles dealing with a more restricted or immediately topical field than those in review journals. (The Editor should be consulted in advance by anyone who considers submitting such a review.)

3. Communications will often describe results obtained by specialized techniques (for example, those of X-ray crystallography and of genetics), and it is most important that technical terms should be avoided or be clearly defined; in any event the actual results obtained and their implications should be described in a way which does not demand from the reader specialized knowledge of a particular technique.

Papers should contain the maximum density of information compatible with clarity, and every sentence should be scrutinized to ensure that its presence is essential. Authors should include enough experimental details to enable others to repeat the work (these may be printed in small type). Careful preparation of manuscript will aid prompt publication; carelessness will cause delay.

4. *Submission of Manuscripts.* Manuscripts should be sent to Dr. J. C. Kendrew, Peterhouse, Cambridge, England, or to any member of the Editorial Board. The submission of manuscript will be taken to imply that the material is original, and that no similar paper has been or will be submitted for publication elsewhere.

Papers will normally be published in English. Manuscripts in French or German will also be accepted, though the use of these languages may cause delay in publication.

5. *Copy.* Papers should be typewritten in double spacing *throughout* (including tables, legends and footnotes), on one side of quarto paper, with a margin of 1 inch all round. The position of tables and illustrations should be indicated in the text; tables and legends for illustrations should be typed separately at the end of the manuscript. All pages should be numbered serially and securely fastened together. Two copies of each manuscript should be submitted, together with three copies of the summary (see §6) and four copies of the abstract (see §9).

A short running title of not more than 40 characters (including spaces), suitable for page headings, should be indicated if the full title is longer than this.

The name of the laboratory where the work has been done should be indicated on the title page. All typescript (including the references) should be carefully checked for errors.

6. *Summary.* A summary will be printed at the head of all papers (excluding Letters to the Editor); this should generally not exceed 300 words, and should be intelligible to the general reader without reference to the main text. Abbreviations should be avoided. Summaries of papers to be printed in French or German will be in English, and should be longer (up to 1000 words) for the benefit of readers not familiar with these languages.

7. *Conventions.* The conventions used in current issues of the Journal for headings, references, etc., should be followed in preparing manuscripts. References should be listed alphabetically at the end of the paper, and abbreviated according to the *World List of Scientific Periodicals*, except that for periodicals with English names capital initial letters should be used throughout. Citations in the text should read thus: Smith, Brown & Robinson (1957), or (Smith, Brown & Robinson, 1957). When a citation has more than two authors, the names of all of them should be given at the first citation (except for more than five authors), and on subsequent citations thus: (Smith *et al.*, 1957). The conventions (Brown, 1957*a*), (Brown, 1957*b*) should be used where more than one paper by the same author(s) has appeared in one year. Volume number should be indicated in the typescript thus: **35**.

Footnotes should be avoided whenever possible. Essential footnotes should be indicated by superscript figures in the text and collected on a single page placed at the end of the manuscript.

Customary abbreviations will be used; authors may refer to *Chemical Abstracts*, to the *Biochemical Journal*, **66**, 8 (1957), and to *Notes on the preparation of papers to be communicated to the Royal Society* (1951). Certain non-standard abbreviations of chemical substances are convenient, and those listed in the *Journal of Biological Chemistry* may be used without further definition (e.g. DNA, ATP). Any other abbreviations of this kind should be used sparingly, and must be defined in a footnote at the point in each manuscript where the first such abbreviation occurs.

8. *Illustrations.* It is often convenient to make a *drawing* two or three times larger than it will appear in the Journal, but very large drawings should be avoided and it will be helpful if drawings and illustrations are submitted on quarto (or 8½ by 11 inch) sheets. Drawings should be made in Indian ink on tracing linen, Bristol board, or faintly blue-lined graph paper; the originals should always be submitted with the manuscript, since photo-prints of drawings often cannot be reproduced satisfactorily. Explanations should, as far as possible, be placed in the legend; the drawing itself should contain a minimum of lettering, and this should be inserted lightly in *blue pencil*, in lower case except for proper names, when a capital should be used. It should be remembered that drawings will be reduced to a size not exceeding 8 by 5 inches, and that the final thickness of the lines should be 0·25 to 0·3 mm. If a curve is given merely for illustrative purposes no cross-lines are needed, but the axes should be graduated at intervals, and the diagram should be boxed in. Where possible, related diagrams should be grouped to form a single figure, and should be prepared uniformly and *on the same scale.* Where appropriate and possible a scale should be included in the illustration itself; the magnification may be noted in the legend and adjusted after reduction by the blockmaker.

Whenever possible *photographs* should be supplied the same size as they are to appear in the Journal. The page area available for a plate is 5 by 7 inches or, at the most, 6 by 8 inches.

Line drawings will be referred to as Figure 1, Figure 2, etc., and half-tone blocks as Plate I, Plate II, etc. The name(s) of the author(s) and the number should be written on all illustrations, together with the indication "Top" if necessary. Besides the original illustrations *one* duplicate set must accompany the second copy of the manuscript; these may be photo-prints, and need not be of the same quality as the originals but must be clear and legible. If *oversize* original drawings are submitted (bigger than 8½ by 11 inches) they must be accompanied by *two* sets of small photo-prints.

9. *Abstracts.* Authors should send with their manuscripts *four copies* (typewritten in double spacing) of an abstract, in about 100 words, suitable for inclusion in biological abstracting journals. This abstract will not appear in the *Journal of Molecular Biology* but will be scrutinized by the Editors before being passed for publication. The abstract should outline as briefly as possible the results and definitive conclusions of the work. Details of methods are generally not required.

The abstract should be headed as follows: title of paper; name(s) of author(s); *J. Mol. Biol.*, followed by space for the year, volume and page references; abbreviated postal address (including country of origin) to which applications for reprints may be sent. For example: The isolation and properties of deoxyribonucleoprotein particles containing single nucleic acid molecules. G. Zubay & P. Doty. J. Mol. Biol. . . . . (Dept. of Chem., Harvard Univ., Cambridge, Mass., U.S.A.). —text of abstract follows.

10. *Reprints.* Fifty reprints of each paper will be supplied free of charge; additional copies may be purchased.

HORMONES IN BLOOD

Edited by

C. H. GRAY, *King's College Hospital Medical School, London*
and A. L. BACHARACH, *Formerly of Glaxo Laboratories Ltd.*

August, 1961, 654 pages, 140s.

CONTENTS:

Introductory
By C. H. GRAY *and* A. L. BACHARACH

Insulin
By P. J. RANDLE *and* K. W. TAYLOR

The Thyroid-Stimulating Hormones and the Iodine-Containing Hormones
By JACOB ROBBINS, J. E. RALL *and* P. G. CONDLIFFE

Oxitocin
By J. M. WALKER

Vasopressin
By J. M. WALKER

Melanotrophin (The Melanocyte-Stimulating Hormone)
By J. M. WALKER

Adrenocorticotrophin
By GEORGE SAYERS

Pituitary Growth and Lactogenic Hormones
By A. KORNER

Pituitary Gonadotrophins
By MICHAEL APOSTOLAKIS *and* JOHN A. LORAINE

Human Chorionic Gonadotrophin (HCG)
By JOHN A. LORAINE

The Oestrogens
By VINCENT J. O'DONNELL *and* JOHN R. K. PREEDY

The Androgens
By ALFRED M. BONGIOVANNI *and* J. DARREL SMITH

Progesterone
By R. V. SHORT

The Corticosteroids
By C. H. GRAY, J. M. GREENAWAY *and* N. H. HOLNESS

The Catecholmines: Adrenaline; Noradrenaline
By U. S. VON EULER

Retrospect and Prospect
By RAYMOND GREENE

Postscript. Immuno-Assay Methods
By R. M. EHRLICH *and* P. J. RANDLE

AUTHOR INDEX—SUBJECT INDEX.

ACADEMIC PRESS, London and New York

Announcing a new serial publication . . .

Advances in

IMMUNOLOGY

Edited by

W. M. TALIAFERRO
Argonne National Laboratories, Illinois

J. H. HUMPHREY
National Institute for Medical Research, London

VOLUME 1 *December* 1961, *about* 450 *pp.*, $12.00 86*s.*

Contents

Delayed Hypersensitivity to Simple Protein Antigens
By P. G. H. GELL, *University of Birmingham, England, and* B. BENACERRAF, *New York University School of Medicine, New York*

The Antigenic Structure of Tumours
By P. A. GORER, *Guy's Hospital, London*

Duration of Immunity in Virus Diseases
By J. H. HALE, *General Hospital, Newcastle upon Tyne, England*

Transplantation Immunity and Tolerance
By M. HAŠEK, A. LENGEROVA, *and* T. HRABA, *Czechoslovak Academy of Sciences, Prague*

Functions of the Complement System
By ABRAHAM G. OSLER, *The Johns Hopkins University School of Medicine and School of Hygiene and Public Health, Baltimore, Maryland*

Immunologic Tolerance of Non-Living Antigens
By RICHARD T. SMITH, *University of Florida College of Medicine, Gainesville*

In Vitro Studies of the Antibody Response
By ABRAM B. STAVITSKY, *Western Reserve University School of Medicine, Cleveland, Ohio*

Fate and Biological Action of Antigen-Antibody Complexes
By WILLIAM O. WEIGLE, *University of Pittsburgh School of Medicine, Pennsylvania*

AUTHOR INDEX—SUBJECT INDEX.

ACADEMIC PRESS, New York and London

Biological Structure and Function

Proceedings of the First IUB/IUBS
International Symposium
Held in Stockholm, September 1960

Edited by

T. W. GOODWIN
Institute of Rural Science, Penglais, Wales

O. LINDBERG
The Wenner-Gren Institute for Experimental Biology, University of Stockholm

THIS important symposium was the first organized by the Biochemistry Section Committee, a coordinating committee between the International Union of Biological Sciences and the International Union of Biochemistry. In accordance with the committee's purpose of choosing subjects in which both biochemistry and the biological sciences were combined in a rapidly expanding field of knowledge, Biological Structure and Function was chosen as the subject of this first symposium. Seventy-two scientists from eleven countries presented invited papers.

VOLUME 1, October 1961, 363 pp., 70s.

Macromolecular Structure and Function
Microsomes and Protein Synthesis
Polysaccharides
Author Index-Subject Index

VOLUME 2, November 1961, about 684 pp., 110s.

Mitochondrial Structure and Function
Structure and Function of Chloroplasts and Chromatophores
Intact Cellular Structure and Function
Specific Membrane Transport and Its Adaptation
Author Index-Subject Index.

ACADEMIC PRESS, *London and New York*

28 MAY 63 F
FILED IN STACKS
13 Feb 63 F
LIBRARY LOAN

CONTENTS

RICH, A. and CRICK, F. H. C., The Molecular Structure of Collagen . . 483
WADA, A., Regularity in the Chain Configuration of D,L-copoly-γ-Benzyl Glutamate 507
WEIS-FOGH, T., Thermodynamic Properties of Resilin, a Rubber-like Protein 520
WILLIAMS, J., CLEGG, J. B. and MUTCH, M. O., Coincidence and Protein Structure 532
GLIMCHER, M. J., BONAR, L. C. and DANIEL, E. J., The Molecular Structure of the Protein Matrix of Bovine Dental Enamel 541
MARVIN, D. A., SPENCER, M., WILKINS, M. H. F. and HAMILTON, L. D., The Molecular Configuration of Deoxyribonucleic Acid. III: X-ray Diffraction Study of the *C* form of the Lithium Salt 547
LUZZATI, V., CESARI, M., SPACH, G., MASSON, F. ET VINCENT, J. M., La Structure du Poly-L-γ-Glutamate de Benzyle en Solution. Configuration en Hélice Différente de l'Hélice α et Transitions entre Formes Hélicoïdales 566
MARMUR, J. and DOTY, P., Thermal Renaturation of Deoxyribonucleic Acids 585
SCHILDKRAUT, C. L., MARMUR, J. and DOTY, P., The Formation of Hybrid DNA Molecules and their use in Studies of DNA Homologies . . 595
WORTHINGTON, C. R., X-Ray Diffraction Studies on the Large-Scale Molecular Structure of Insect Muscle 618
LUZZATI, V., MASSON, F. and LERMAN, L. S., Interaction of DNA and Proflavine: A Small-Angle X-Ray Scattering Study 634
FRASER, R. D. B. and MACRAE, T. P., The Molecular Configuration of α-Keratin 640
WEIS-FOGH, T., Molecular Interpretation of the Elasticity of Resilin, a Rubber-like Protein 648
CHUN, E. H. L. and LITTLEFIELD, J. W., The Separation of the Light and Heavy Strands of Bromouracil-substituted Mammalian DNA . . . 668
LEVINTHAL, C. and DAVISON, P. F., Degradation of Deoxyribonucleic Acid under Hydrodynamic Shearing Forces 674
NOMURA, M. and BENZER, S., The Nature of the "Deletion" Mutants in the *r*II Region of Phage T4 684

Letters to the Editor—

The Effect of 8-Azaguanine on the Ribonucleic Acid of the Ribosomes in *Bacillus cereus*, E. OTAKA, S. OSAWA and Y. OOTA . . . 693
The Electron Microscopy of Phage DNA Molecules with Denatured Regions, M. BEER and C. A. THOMAS, JR. 699
Surface Structure of Tobacco Mosaic Virus, R. G. HART 701
A Temperature-Sensitive Regulatory System, T. HORIUCHI, S. HORIUCHI and A. NOVICK 703
Sidechain-Sidechain Interaction between the α-Helices of L- and D-Polypeptide Chains, M. TSUBOI, A. WADA and N. NAGASHIMA . . 705
Haemoglobin I_2, a Further Variant of Haemoglobin A_2, J. ATWATER, E. R. HUEHNS and E. M. SHOOTER 707
The Reaction of Hydroxylamine with Cytosine and Related Compounds, D. M. BROWN and P. SCHELL 709

Printed in Great Britain at the Aberdeen University Press